FOR FURTHER TRAN ~~~~ ②

BASIC OPTIMAL ESTIMATION AND CONTROL PROBLEMS IN HILBERT SPACE

by

R. M. DeSantis, R. Saeks, and L. Tung

1978

D D C

JUN 23 1978

E
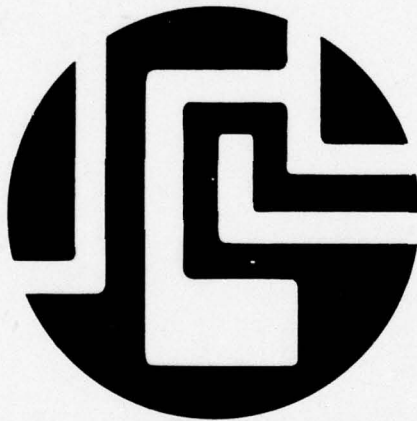
Institute for
Electronics Science

TEXAS TECH UNIVERSITY
Lubbock, Texas 79409

78 06 19 098

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR—78-1046	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) BASIC OPTIMAL ESTIMATION AND CONTROL PROBLEMS IN HILBERT SPACE		5. TYPE OF REPORT & PERIOD COVERED Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R.M. DeSantis, R. Saeks and L. Tung		8. CONTRACT OR GRANT NUMBER(s) AFOSR-74-2631
9. PERFORMING ORGANIZATION NAME AND ADDRESS Texas Technological University Department of Electrical Engineering Lubbock, Texas 79409		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A6
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE 1978
		13. NUMBER OF PAGES 72
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) 76 p.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Resolution Space   Kalman Filter
Wiener-Hopf Optimization
Control
Estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The recently developed mathematical framework of Hilbert resolution space valued random processes is used to formulate and solve an abstract quadratic optimization problem. By particularizing the description of the operators appearing in the statement and solution formula of this problem one rediscovers and generalizes most of the classical estimation and control theory problem statements and results. These results include, among others, the Wiener smoothing prediction filter, the Kalman regulator, the

DD FORM 1 JAN 73 1473

410 354

20. Abstract

Kalman-Bucy filter, the stochastic control separation principle and the more recent Youla-Jabr-Bongiorno optimal servo problem solution.

BASIC OPTIMAL ESTIMATION AND CONTROL PROBLEMS

IN HILBERT SPACE[*]

De Santis[#], R.M., Saeks[+], R., Tung[o], L.J.

0. ABSTRACT

The recently developed mathematical framework of Hilbert
resolution space valued random processes is used to formulate
and solve an abstract quadratic optimization problem. By parti-
cularizing the description of the operators appearing in the
statement and solution formula of this problem one rediscovers
and generalizes most of the classical estimation and control
theory problem statements and results. These results include,
among others, the Wiener smoothing prediction filter, the Kalman
regulator, the Kalman-Bucy filter, the stochastic control separation
principle and the more recent Youla-Jabr-Bongiorno optimal servo
problem solution.

---

o Department of Electrical Engineering, University of Texas, at
  El Paso, El Paso, Texas 79968

Ecole Polytechnique de Montréal, University of Montreal, Montreal.

+ Texas Tech University, Lubbock, Texas.

TABLE OF CONTENTS

LIST OF SYMBOLS

Symbol	Meaning	Section first used
HRS	Hilbert Resolution Space	-
P. P^t, P_t	orthoprojector in a HRS	2
H	Hilbert space	2
x, y, z	elements of H	2
$|x|$	the norm of x	2
R	resolution of the identity in H	2
ν	linearly ordered set (usually the real numbers)	2
t_o, t_∞	minimum and maximum elements of ν	2
t, k, ℓ, i	elements of ν	2
$[H, P^t]$	Hilbert Resolution Space with $R = \{P^t, t\epsilon\nu\}$	2
$(H, \Sigma, \mathcal{P})$	a space of Hilbert space valued random processes	2
Σ	family of Borel sets in H	2
$\mathcal{P}$	probability measure on Σ	2
ρ, β, α, η, π	elements of $(H, \Sigma, \mathcal{P})$	2
Q_ρ	the covariance operator associated with the random process ρ	2
$Q_{\alpha\beta}$	the cross covariance operator associated with the random processes α and β	2
m_ρ	the mean of the random process ρ	2
$E\left[f(\rho)\right]$	the expected value of the random variable $f(\rho)$	2
L, T, N, D	operators acting on H	2

iv

BASIC OPTIMAL ESTIMATION AND CONTROL PROBLEMS
IN HILBERT SPACE[#]

1. INTRODUCTION

Most of the classical [Wi 1] and modern [Yo 1], Wiener-
Kolmogoroff linear least square optimization methodology is
based upon frequency domain and analytic function theory proce-
dures. As a consequence, most of its applications traditionally
focus on the realm of stationary, lumped parameters and infinite
time interval systems [Ne 1], [Ch 1]. As Bode and Shannon [Bo 1]
observed, however, the computational steps involved in these
procedures have a direct physical interpretation in terms of
causality related operations and random signal representations
whose meaningfulness does not depend on any one of these special
properties. Thus, it is perfectly natural to look for extensions
of the Wiener-Kolmogoroff methodology so as to make it applicable
to systems of a more general type. More specifically, the following
questions are of interest: can the Wiener-Kolmogoroff methodology
generalize to a form which is independent of analytic function theory?
can this potential generalization encompass multivariate cases and
related frequency domain results which utilize the algebraic Riccati
equation? can nonstationary systems, finite time interval, and in-
finite dimensional state space problems be solved using the genera-
lized solution?

By De Santis, R.M., Saeks, R., and Tung, J.L.

1

The objective of the present paper is to show that, on the
basis of recent developments about causality [De 1], [De 4],
[Sa 1], [Sa 2], state [Sa 6], [Sch 1], [St 1] and stochastic
signals representation [Kai 1], [Kai 2], [Sa 3], [Ba 2], the
answer to the above questions is an affirmative one. It is
indeed now possible to develop a generalized Wiener-Kolmogoroff
methodology which while encompassing classical optimal estimation
and control results is also applicable to nonstationary, finite
time interval, and distributed parameters systems. The advantages
of such a generalization are multiple: a better perspective of the
relations among a number of different problems and techniques; a
clarification of the role played by causality, state and related
system theoretic concepts; a wider range of application of the
available results.

We do this by combining the classical Wiener-Kolmogoroff
ideas and techniques [Wi] with the novel framework of abstract
Hilbert resolution spaces [Go 2]. We start by formulating and
solving two simple optimization problems, (statements P1 and P2,
Theorems 1 and 2). Subsequently we show that by appropriately
specializing the description of the operators appearing in the
statements and related solution formulas of these problems, one can
rediscover and generalize the solutions of most of the optimal esti-
mation and control problem of interest, namely: the Wiener smoothing
and prediction filter, (Theorem 3), the Youla-Jabr-Bongiorno optimal
servo compensator, (Theorem 4), the Porter basic optimization

formula, (Theorem 5), the Kalman optimal regulator and the principle
of optimality, (Theorem 6), the Kalman-Bucy filter, (Theorem 7), and
the separation principle, (Theorem 8).

The development inherits much from the previous work on the
subject. The questions under consideration were first formulated
and studied by Porter [Po 1]. The mathematical setting is based on
the Nonself-adjoint Volerra operators techniques developed by
Gohberg and Krein [Go 1], [Go 2] and on the more recent Hilbert space
valued random processes framework proposed by Balakrishnan [Ba 1].
In regard to the results, the solution of our basic optimization
problem uses ideas and techniques from Porter [Po 1], Bode and Shannon
[Bo 1], Kailath [Kai 1], Balakrishnan [Ba 1], and Youla [Yo 1]. A
significant role in the proofs of the state estimation and control
results has been played by the abstract state space realization work
by Steinberger, Schumitzky and Silverman [St 1].

2. MATHEMATICAL BACKGROUND

The reader will be assumed to be familiar with those standard notions which are usually associated with the concepts of Banach and Hilbert space [Po 2, chp 1], [Ba 2, chp 1]. In addition to these we will also use the concepts of Hilbert resolution space which are discussed in [Go 1, chp 1], [Sa 2], plus the concepts of Hilbert space valued random variables [Ba 2, chp 6], [Sa 3]. These latter concepts will be briefly reviewed in the following paragraphs.

A bounded linear operator P on a Hilbert space H is called an __orthoprojector__ if for all pairs x, yϵH one has that $\langle Px, y \rangle = \langle x, Py \rangle$ and $P(Px) = Px$. Given a linearly ordered set ν with minimum and maximum elements t_0 and t_∞, we will say that a family of orthoprojectors, $R = \{P^t: t\epsilon\nu\}$, is a __resolution of the identity__ if it enjoys the following two properties:

i) $P^{t_0}H = 0$, $P^{t_\infty}H = H$ and $P^k H \supseteq P^\ell H$ whenever
   $k > \ell$

ii) if $\{P^i\}$ is a sequence of orthoprojectors in R and there
   exists an orthoprojector P such that $\{P^i x\} \to Px$ for each
   xϵH, then PϵR.

A Hilbert space H equipped with a resolution of the identity $R = \{P^t: t\epsilon\nu\}$ is called a __Hilbert resolution space__ and is denoted by the symbol $[H, P^t]$. Given an orthoprojector $P^t\epsilon R$, the orthoprojector $I-P^t$ is often denoted by the symbol P_t.

4

We will use the triple $(H, \Sigma, \mathcal{P})$ to denote a probability space of <u>Hilbert space valued random processes</u> with Hilbert space, H, family of Borel sets, Σ, and probability measure, $\mathcal{P}$. Recall that [Ba 2, chp 6], with each random process, $\rho \epsilon (H, \Sigma, \mathcal{P})$, with a finite first and second moment, one finds associated an element, $m_\rho \epsilon H$, (the <u>mean</u> value of ρ) and an operator, $Q_\rho : H \rightarrow H$, (the <u>covariance</u> of ρ), such that:

$$E(x, \rho) = (x, m_\rho), \forall x \epsilon H$$

and $\qquad E(x, \rho - m_\rho) (y, \rho - m_\rho) = (x, Q_\rho), \forall x, y \epsilon H$

where the symbol "E $f(\rho)$" denotes the expected value of the scalar valued random variable $f(\rho)$ with respect to the probability space underlying ρ. Similarly, with each pair ρ, $\pi \epsilon (H, \Sigma, \mathcal{P})$ one finds associated an operator, $Q_{\rho\pi}$, (the <u>cross covariance</u>) with the property that:

$$E(x, \rho - m_\rho) (y, \pi - m_\pi) = (x, Q_{\rho\pi} y), \forall x, y \epsilon H.$$

In the sequel we will not explicitly use the Σ family of Borel sets or the specific nature of the probability measure $\mathcal{P}$. Thus, for more background about these concepts we simply refer the reader to [Ba 2, chp 2].

The mean values and the covariance and cross covariance operators, however, will be used extensively. While the theory associated with these concepts is again well documented in [Ba 2, chp 6], for our purposes it will be helpful to list a few of their most important properties namely:

(i) $\quad m_{(\rho+\pi)} = m_\rho + m_\pi$;

(ii) $|m_\rho| \leq E \{|\rho|\}$;

(iii) $m_{L\rho} = Lm_\rho$ for any bounded linear operator L;

(iv) $Q_{(L\rho)}{}_{(T\pi)} = LQ_{\rho\pi} T^*$ for bounded linear operators L and T;

(v) $Q_{\rho+\pi} = Q_\rho + Q_{\rho\pi} + Q_{\pi\rho} + Q_\pi$;

(vi) $Q_{\rho\pi} = Q_{\pi\rho}^*$;

(vii) $E \{|\rho|^2\}$ is finite if and only if Q_ρ is nuclear in which
case $E \{|\rho|^2\} = tr(Q_\rho)$;

(viii) If $E \{|\rho|^2\}$ is finite, then $Q_{\rho\pi}$ is Hilbert-Schmidt for
all random processes π with a finite second moment.

Note that given an operator, $T: H \to H$, the trace of T is defined by:

$$tr(T) = \sum_i (Te_i, e_i)$$

where $\{e_i\}$ is any orthonormal basis in H. A useful property of the
trace is that for T_1, $T_2: H \to H$ one has

$$tr(T_1 T_2) = tr(T_2 T_1).$$

The random processes ρ and $\pi\epsilon$ (H, Σ, $\mathcal{P}$) will be said to be <u>sta-
tistically independent</u> if $Q_{\rho\pi} = 0$. In the course of our development
we will use the fact that linear transformations of independent random
processes are independent, plus the fact that if ρ and π are indepen-
dent then $Q_{\rho+\pi} = Q_\rho + Q_\pi$. Unless otherwise specified we will only
consider zero mean valued random processes.

3. CAUSALITY AND STATE CONCEPTS

The concepts of causality and state will play a fundamental role in the development of the following sections. While a full treatment of this subject can be found in a number of references, (among others [De 1], [Sa 2], [Sa 6], [Sch 1], and [St 1]), it is helpful to review some of the basic definitions and results.

3.1 Basic Definitions and Properties

An operator T: $[H, P^t] \to [H, P^t]$ is said to be __causal__ if $P^t x = P^t y$ implies $P^t T x = P^t T y$. T is __strictly causal__ if it is causal and for any given $\varepsilon > 0$ one can find a partition $\{t_o = \xi_o, \xi_1, \ldots, \xi_n = t_\infty\}$ $\varepsilon \nu$ such that

$$\sup_i |\Delta_i T \Delta_i| < \varepsilon$$

where $\Delta_i = P^{\xi_i} P_{\xi_{i-1}}$. T is called __anticausal__ (strictly anticausal) if T^* is causal (strictly causal). T is __memoryless__ if it is simultaneously causal and anticausal. The validity of the following results is easily verifiable.

__Lemma 1:__ The following statements are equivalent: T is causal; $P^t x = 0$ implies $P^t T x = 0$, $T P_t = P_t T P_t$; $P^t T = P^t T P^t$; T^* is anticausal.

__Lemma 2:__ Suppose that T_1 is causal and T_2 is strictly causal. Then $T_1 T_2$ and $T_2 T_1$ are strictly causal.

Our definition of strict causality even though most convenient from a technical point of view is in general more restrictive than what it will be really needed. Indeed, in most of our results the requirement that $\sup_i |\Delta_i T \Delta_i| < \varepsilon$ could be replaced by $\sup_i |\Delta_i T \Delta_i x| < \varepsilon$, where the partition

7

is chosen in correspondence with a given (x, ϵ) pair. A notable exception to this is given by the following lemma whose validity requires the restrictive form of our definition.

Lemma 3: Suppose that T is a strictly causal operator. Then $I + T$ is invertable with $(I + T)^{-1} = I + K$, where K is strictly causal and such that $K = \sum_n (-1)^n T^n$.

Given a strictly causal, (strictly anticausal), $T: [H, P^t] \rightarrow [H, P^t]$, a *state realization*, (costate realization), of T is given by a triplet $(\Psi_T(t), H_T(t), \xi_T(t))$ such that for each $t \epsilon \nu$ one has that

$H_T(t)$ is a Hilbert space,

$\Psi_T(t): P^t H \rightarrow H_T(t)$, $(P_t H \rightarrow H_T(t))$,

$\xi_T(t): H_T(t) \rightarrow P_t H$, $(H_T(t) \rightarrow P_t H)$,

and, for each $u \epsilon H$, $\xi_T(t) \Psi_T(t) u = P_t T P^t u$, $(= P^t T P_t u)$. The term state realization is suggested by the fact that for every $t \epsilon \nu$ one has that

$$P_t T u = \xi_T(t) x(t) + P_t T P_t u$$

where $x(t) = \Psi_T(t) u$. Thus, the element $x(t)$, the *state* of T at t, possesses all the information which is necessary to determine the influence of the past values of the input over the future values of the output. In [Sch 1] it is shown that every strictly causal operator admits a state realization.

Given a strictly causal, (strictly anticausal), $T: [H, P^t] \rightarrow [H, P^t]$, with state realization, (costate realization), $(\Psi_T(t), H_T(t), \xi_T(t))$, the pair (k_T, g_T) is called a *state trajectory realization*, (*costate trajectory realization*), if $T = g_T k_T$ and

$$k_T: [H, P^t] \to X_T, \quad k_T \text{ strictly causal, (strictly anticausal)};$$

$$g_T: X_T \to [H, P^t], \quad g_T \text{ memoryless, and}$$

$$X_T: \left[L_2 [\nu, H_T], P^t \right], \text{ where } H_T = \operatorname*{Span}_{t \in \nu} H_T(t).$$

Here the term state trajectory is justified by the fact that the element $x = k_T u$, the <u>state trajectory</u> associated with u, can be viewed as a mapping, $x(\cdot): \nu \to H_T(t)$, where $x(t) = \Psi_T(t)u$ represents the state of T at t. The <u>trajectory space</u> X_T is built from ν and H_T by following the procedure outlined in [Ba 2, section 3.5]; in doing this the linearly ordered set ν is tacitly assumed to be a measurable subset of the real line. For an exhaustive discussion about this concept the reader is addressed to [St 1]. Among other things, this reference proves that every strictly causal T has a state trajectory realization.

In working with the above causality and state concepts it will be useful to make use of the alphabetic code $A = \{A, C, M, \bar{A}, \bar{C}, Co, St\}$ which is defined as follows

A = anticausal	$\bar{C}$ = strictly causal
C = causal	Co = costate
M = memoryless	St = state
$\bar{A}$ = strictly anticausal	

A first application of this code is given by the following <u>principle of causal duality</u>.

Lemma 4: Let a statement or equality by phrased using relations involving concepts associated with the alphabet $A = \{A, C, M, \bar{A}, \bar{C}, St, Co\}$ and the family of projection operators $\{P^t\}$ and $\{P_t\}$. Then the statement or equality remains valid if the following interchange in symbols occurs:

$$P^t \rightarrow P_t, \qquad P_t \rightarrow P^t, \qquad \bar{C} \rightarrow \bar{A}, \qquad \bar{A} \rightarrow \bar{C},$$
$$C \rightarrow A, \qquad A \rightarrow C, \qquad St \rightarrow Co, \qquad Co \rightarrow St.$$

3.2 Causality Additive and Multiplicative Decompositions

An operator T is said to admit a causality additive decomposition [De 1], [Sa 2] if T can be viewed as given by the sum of a causal plus an anticausal component; T is said to admit a canonical causality decomposition if it can be represented in terms of the sum of a strictly causal, a strictly anticausal and a memoryless component. The following lemma establishes the uniqueness of such a decomposition plus a useful property relating the causality decomposition of T with those of FT and TF, where F is a memoryless system.

Lemma 5: Suppose that the operator T has a canonical causality (additive) decomposition, $T = T_{\bar{A}} + T_{\bar{C}} + T_M$. Then the elements T_α, $\alpha\epsilon \{\bar{A}, \bar{C}, M\}$ are uniquely defined. Moreover, if F is memoryless and $T = T_{\bar{A}} + T_{\bar{C}} + T_M$, then

$T_1 = TF$ and $T_2 = FT$ are such that

$T_i = T_{i\bar{A}} + T_{i\bar{C}} + T_{iM}$, $i = 1, 2$

where $T_{1\alpha} = T_\alpha F$ and $T_{2\alpha} = FT_\alpha$, $\alpha\epsilon \{\bar{A}, \bar{C}, M\}$.

In view of lemma 5 , in general it makes sense to talk about the causal (strictly causal, anticausal, ...) component of an operator $T: [H, P^t] \rightarrow [H, P^t]$. For example, given $T = T_{\bar{A}} + T_{\bar{C}} + T_M$, the causal component of T, usually denoted by T_C or $[T]_C$ is given by $T_{\bar{C}} + T_M$. The causal component of an operator T, (as well as the other strictly causal, memoryless, ... etc., components) does not always exist. General necessary and sufficient existence conditions plus a number of interesting causality decomposition related properties can be found in [De 1], [Sa 2] and [De 3]. One of these properties will be of a particular interest later on in the development. To characterize it, consider a partition $\Omega = \{\xi_o = t_o, \xi_1, \ldots, \xi_n = t_\infty\} \epsilon\nu$, and associate with it the operator valued function

$$\Phi(\Omega) = \sum_{i=1}^{n} \Delta_i TP^i$$

where $\Delta_i = P^{\xi i} - P^{\xi i-1}$; $P^i = P^{\xi i}$. If $[T]_C$ is well defined then given any $\epsilon > 0$ one can find a partition Ω_ϵ such that for every $\tilde{\Omega} \supset \Omega_\epsilon$ one has that

$$| [T]_C - \Phi(\Omega) | < \epsilon$$

where the norm can be intended either in the uniform or in the strong operator topology. This property can be expressed using the following symbols

$$\{\Sigma \Delta_i T P^i\} \rightarrow (M) \int dPTP^S \rightarrow [T]_C.$$

In the particular case of Hilbert-Schmidt operators a canonical causality decomposition is always well defined. Moreover the components of this decomposition are mutually orthogonal. All this is established by the following lemma.

Lemma 6: Causality and Hilbert-Schmidt operators

i) the class of causal, (anticausal, memoryless), Hilbert-Schmidt, operators is a Hilbert space;

ii) in the Hilbert-Schmidt inner product topology one has

$< T_\alpha, T_\beta > = 0$ whenever $\alpha \neq \beta \epsilon \{\bar{A}, \bar{C}, M\}$;

iii) if T is a Hilbert-Schmidt operator then

$$T = T_{\bar{A}} + T_{\bar{C}} + T_M$$

where the T_α, $\alpha \epsilon \{\bar{A}, \bar{C}, M\}$, are well defined Hilbert-Schmidt operators such that

$$||T - T_\alpha|| = \frac{\min}{\tilde{T}_\alpha} ||T - \tilde{T}_\alpha||$$

where $\tilde{T}_\alpha$ is any other α operator and the symbol $|| \ ||$ denotes the Hilbert-Schmidt norm.

Equal in importance to the concept of a causality additive decomposition is that of causality multiplicative decomposition also referred to as causal factorization. An operator $Q: [H, P^t] \rightarrow [H, P^t]$ admits a causal factorization if Q can be represented in terms of the cascade composition of a causal (anticausal) operator followed by an anticausal (causal) component. In the sequel we will mainly be concerned with operators Q which are positive Hermitian ($<Qx, x> \geq 0$, $Q = Q^*$); and with causal factorizations of the type considered in the following lemma [Sa 3].

Lemma 7: If $Q: [H, P^t] \to [H, P^t]$ is positive Hermitian then

i) there exists a causal and causally left invertable one-to-one

$\Omega: [H, P^t] \to [\bar{H}, \tilde{P}^t]$ such that

$$Q = \Omega\Omega^*$$

ii) there exists a causal and causally right invertable one-to-one

$\Lambda: [H, P^t] \to [\bar{H}, \tilde{P}^t]$ such that

$$Q = \Lambda^*\Lambda$$

3.3 Causal Decomposition and State Realization

The following two lemmas generalize two well known linear time invariant results ([Bar 1], [Ma 1]). They unveil some key relations concerning a special multiplicative causality decomposition, the causal part of a related operator and the concept of state trajectory. It is precisely by virtue of these relations that we will be able to elucidate the connections between a plant causality and state structure on the one hand, and the structure of the optimal filter and/or compensator on the other. The proof of the first part of these lemmas can be found in [St 1]. The second part follows by involving the principle of causal duality. In stating these lemmas, the symbols F and B will denote two memoryless operators, while G will represent a strictly causal system. We will suppose that (GB, F) is a state trajectory realization of FGB and (G*F*, B*) is a costate trajectory realization of $B^*G^*F^*$.

<u>Lemma 8</u>:

    i) Suppose that there exists a strictly causal V_1 such that

$$I + B^*G^*F^*FGB = (I + V_1^*)(I + V_1)$$

then (†)

$$\left[(I + V_1^*)^{-1} B^*G^*F^*FGB\right]_C = V_1.$$

    ii) If there exists a strictly causal V_2 such that

$$I + FGBB^*G^*F^* = (I + V_2)(I + V_2^*)$$

then

$$\left[FGBB^*G^*F^*(I + V_2^*)^{-1}\right]_C = V_2.$$

<u>Lemma 9</u>:

    i) Suppose that there exists a strictly causal V_1 such that

$$I + B^*G^*F^*FGB = (I + V_1^*)(I + V_1)$$

then there exists a well defined memoryless M_R such that

$$V_1 = M_R GB.$$

    ii) If there exists a strictly causal V_2 such that

$$I + FGBB^*G^*F^* = (I + V_2)(I + V_2^*)$$

then there exists a well defined memoryless M_e such that

$$V_2^* = M_e^*G^*F^*.$$

    The next lemma provides the key to the stochastic control principle of separation (section 6).

(†) Recall that the symbol $\left[T\right]_C$ denotes the causal component of T.

Lemma 10: Suppose that there exists strictly causal V_1 and V_2 such that

$$I + B^* G^* F^* FGB = (I + V_1^*) (I + V_1)$$

and

$$I + FGBB^* F^* = (I + V_2) (I + V_2^*)$$

Then one has that

$$\left[(I + V_1^*)^{-1} B^* G^* F^* FGBB^* G^* F^* (I + V_2^*)^{-1} \right]_C = M_R G M_e$$

where M_R and M_e are memoryless and such that

$$V_1 = M_R GB \quad \text{and} \quad V_2 = FGM_e.$$

Proof: From the strict anticausality of $G^* F^* (I + V_2^*)^{-1}$ one has

$$\left[(I + V_1^*)^{-1} B^* G^* F^* FGBB^* G^* F^* (I + V_2^*)^{-1} \right]_C =$$

$$\left[\left[(I + V_1^*)^{-1} B^* G^* F^* FGB \right]_C B^* G^* F^* (I + V_2^*)^{-1} \right]_C$$

Moreover, from lemma 8i)

$$\left[(I + V_1^*)^{-1} B^* G^* F^* FGB \right]_C = V_1$$

and, from lemma 8ii)

$$\left[FGBB^* G^* F^* (I + V_2^*)^{-1} \right]_C = V_2$$

Using these equalities and applying lemma 9, it follows that

$$\left[(I + V_1^*)^{-1} B^* G^* F^* FGB \right]_C = M_R GB$$

and

$$\left[FGBB^* G^* F^* (I + V_2^*)^{-1} \right]_C = FGM_e.$$

This leads to

$$\left[\left[(I + V_1^*)^{-1} B^* G^* F^* FGB\right]_C B^* G^* F^* (I + V_2^*)^{-1}\right]_C =$$

$$M_R F^{-1} \left[FGBB^* G^* F^* (I + V_2^*)^{-1}\right]_C = M_R GM_e.$$

where F^{-1} has been used to denote the pseudo inverse of F.

4. TWO FUNDAMENTAL PROBLEMS

In the course of the development it will be shown that many of the quadratic optimal control problems considered in the technical literature can be viewed as special cases of a "basic stochastic optimal control problem"; a key step in the solution of this problem will in turn be provided by the solution of what we will call a "best causal approximant problem". With this in mind we find it convenient to proceed as follows. We start by stating and solving the best causal approximant problem. Subsequently we discuss the basic stochastic optimal control problem. In the forth coming sections we will show how these preliminary results allow us to solve a variety of other problems of interest.

4.1 The Best Causal Approximant Problem

The best causal approximant problem can be stated as follows (see Figure 1).

<u>Statement P1</u>: The best causal approximant problem.

Given a not necessarily causal $\Pi: [H, P^t] \to [H, P^t]$, and a random process $z \epsilon (H, \Sigma, P)$, with covariance operator Q_z, determine a causal Φ_o such that for any other causal Φ one has that

$$J(\Phi_o) \leq J(\Phi)$$

where

$$J(\Phi) = E \{ ||(\Phi - \Pi)z||^2 \} = tr \{ (\Phi - \Pi)Q_z (\Phi - \Pi)^* \}.$$

The following theorem gives us a sufficient condition for the solution of the best causal approximant problem.

17

__Theorem 1__: A sufficient condition for a causal Φ_0 to be a solution of the best causal approximant problem is that

$$\Omega\Phi_0 = [\Pi\Omega]_C$$

where Ω is causal and such that

$$Q_z = \Omega\Omega^*.$$

__Proof__: Let us first consider the case where $Q_z = I$, that is the case where z is a white noise random process. For any partition $\{\xi_0 = t_o, \xi_1, \ldots, \xi_N = t_\infty\}$ $\epsilon\nu$ one has

$$J(\Phi) = E\{|(\Pi - \Phi)z|^2\}$$

$$= \sum_{i=1}^{N} E\{|\Delta_i (\Pi P^i - \Phi)z + \Delta_i \Pi P_i z|^2\}$$

where $P^i = P^{\xi_i}$ and $\Delta_i = P^i P_{i-1}$. From here, taking into account the statistical independence of $P^i z$ and $P_i z$, it follows

$$J(\Phi) = \sum_{i=1}^{N} E\{|\Delta_i (\Pi P^i - \Phi)z|^2\} + \sum_{i=1}^{N} E\{|\Delta_i \Pi P_i z|^2\}$$

Recalling that by a partition refining process we have that $\sum_{i=1}^{N} \Delta_i \Pi P^i$ converges to $[\Pi]_C$, it follows

$$J(\Phi) = E\{|([\Pi]_C - \Phi)z|^2\} + E\{|(\Pi - [\Pi]_C)z|^2\}$$

Since the second term of this sum is independent from Φ, it follows that the problem of minimizing $J(\Phi)$ is equivalent to that of minimizing the first term. This is obviously done by choosing $\Phi_0 = [\Pi]_C$.

Let us consider the more general case where $Q_z = \Omega\Omega^* \neq I$, that is the case where z is a colored random process. This case can be reduced to the special white noise case by regarding z as given by

the output of the causal filter Ω driven at the input by a white noise random process ω. Using the notations $\Pi' = \Pi\Omega$ and $\Phi' = \Phi\Omega$ this leads to

$$J(\Phi) = E \{|(\Pi - \Phi)\Omega\omega|^2\}$$
$$= E \{|(\Pi' - \Phi')\omega|^2\}$$

From here one can repeat the white noise argument to conclude that $J(\Phi)$ is indeed minimized by $\Phi' = \Phi\Omega = [\Pi\Omega]_C$.

Note that Theorem 1 does not imply that $J(\Phi_o) < \infty$. If $J(\Phi)$ is not finite, however, no other Φ will make $J(\Phi)$ finite. These difficulties may be alleviated by assuming that certain operators are Hilbert-Schmidt. Using these assumptions Theorem 1 can also be proved in a more direct and mathematical more transparent fashion. As this alternative fashion puts in a better perspective the role played by Hilbert-Schmidt operators and their special properties it is in order to introduce the following alternative proof.

<u>An alternative proof of Theorem 1</u>: Assume that Φ and Π are Hilbert-Schmidt and observe that

$$J(\Phi) = E \{|(\Phi - \Pi)z|^2\} = \text{tr} \{(\Phi - \Pi)Q_z (\Phi - \Pi)^*\}$$

From $Q_z = \Omega\Omega^*$, with Ω causal, and denoting by $\{e_i\}$, a complete orthonormal set in H, it follows that

$$J(\Phi) = \text{tr} \{(\Phi - \Pi)\Omega\Omega^* (\Phi - \Pi)^*\}$$
$$= \{\underset{i}{\Sigma} < (\Phi - \Pi)\Omega e_i, (\Phi - \Pi)\Omega e_i >\}$$
$$= ||(\Phi - \Pi)\Omega||^2$$

where the symbol $|| \ ||$ denotes the Hilbert-Schmidt norm. Observing

now that $\Phi\Omega$ is causal, and recalling that the causal and strictly anticausal parts of a Hilbert-Schmidt operator are well defined and mutually orthogonal, (lemma 6), it follows that

$$J(\Phi) = \left|\left| \Phi\Omega - [\Pi\Omega]_C \right|\right|^2 + \left|\left| [\Pi\Omega]_{\bar{A}} \right|\right|^2 .$$

From here, observing that $\left|\left| [\Pi\Omega]_{\bar{A}} \right|\right|^2$ is independent of Φ, one can conclude that a sufficient condition for Φ_0 to be the desired solution is

$$\Phi_0\Omega = [\Pi\Omega]_C .$$

The following corollary considers an important special case of the best causal approximant problem.

<u>Corollary 1</u>: Suppose that for a given $P^0 \epsilon R$ one has that

$$Q_z = P_0 G (I + P^0 Q_\omega P^0) G^* P_0$$

with G causal. Then a sufficient condition for Φ_0 to be a solution of the best causal approximant problem is

$$\Phi_0 P_0 G = [\Pi P_0 G]_C .$$

<u>Proof</u>: Observe that in this particular case one has

$$J(\Phi) = J_1(\Phi_1) + J_2(\Phi_2)$$

with

$$J_1(\Phi_1) = \left|\left| (\Phi_1 - \Pi) P_0 G \right|\right|^2$$

and

$$J_2(\Phi_2) = \text{tr} \{ (\Phi_2 - \Pi) P_0 GP^0 Q_\omega P^0 G^* P_0 (\Phi_2 - \Pi)^* \}.$$

The proof can then be completed by verifying that a sufficient condition for these two latter functionals to be simultaneously minimized is

$$\Phi_{o1}G = \Phi_{o2}G = [\Pi P_o G]_C .$$

This is done by observing that

$$J_1(\Phi_1) = ||P_o (\Phi_1 P_o G - [\Pi P_o G]_C)||^2 + ||P^o [\Pi P_o G]_C||^2 + || [\Pi P_o G]_{\bar{A}}||^2$$

and

$$J_2(\Phi_2) = \mathrm{tr} \{P_o (\Phi_2 P_o G - [\Pi P_o G]_C) P^o Q_\omega P^o G^* P_o (\Phi_2 - P_o \Pi)^*$$
$$+ \mathrm{tr} \{P^o \Pi P_o G P^o Q_\omega P^o G^* P_o \Pi^* P^o\} .$$

The last equality follows from

$$P_o \Pi P_o G P^o = P_o [\Pi P_o G]_C P^o$$

and the fact that

$$\mathrm{tr} \{P^o \Pi P_o G P^o Q_\omega P^o G^* P_o (\Phi - P_o \Pi)^* P_o\} = 0$$
$$\mathrm{tr} \{P_o (\Phi_2 - P_o \Pi) P_o G P^o Q_\omega P^o G^* P_o \Pi^* P^o\} = 0 .$$

4.2 The Basic Stochastic Optimization Problem

The following problem is inspired from the "basic" deterministic problem studied by Porter in [Po 1]. In the spirit of this reference we will retain the "basic" qualifier because as we shall see in the later sections, this problem includes as a special case many other fundamental problems such as the Wiener filter, the Kalman optimal regulator, the optimal Kalman-Bucy state estimation problem.

<u>Statement P2</u>:   The basic stochastic optimization problem (Figure 2).
Given four not necessarily causal systems L_i: $[H, P^t] \rightarrow [H, P^t]$,
$i = 1, 2, 3, 4$; and two random processes $\alpha, \beta \epsilon (H, \Sigma, \mathcal{P})$; determine
a causal controller, D_o, such that for any other causal D one has

$$J(D_o) \leq J(D)$$

where

$$J(D) = E \{|e|^2 + |r|^2\}$$

and

$$e = L_3 D (L_1 \alpha + \beta) - L_2 \alpha$$

$$r = L_4 D (L_1 \alpha + \beta).$$

The next theorem, the key result of the overall development, says
that under some appropriate hypotheses the basic optimization problem
is equivalent to an associated best causal approximant problem; its
solution can then be obtained by using Theorem 1.

<u>Theorem 2</u>:   Suppose that there exists causal and causally invertible
Ω and Λ such that

$$Q = L_1 Q_\alpha L_1^* + Q_\beta + L_1 Q_{\alpha\beta} + Q_{\beta\alpha} L_1^* = \Omega\Omega^*$$

and

$$L_3^* L_3 + L_4^* L_4 = \Lambda^* \Lambda .$$

Then the basic optimization problem is equivalent to a best causal
approximant problem with

$$Q_z = Q$$

and

$$\Pi = \Lambda^{*-1} L_3^* (L_2 Q_\alpha L_1^* + L_2 Q_{\alpha\beta}) Q^{-1}.$$

A solution of the basic optimization problem is given by

$$D_o = \Lambda^{-1} \left[\Lambda^{*-1} L_3^* (L_2 Q_\alpha L_1^* + L_2 Q_{\alpha\beta}) \Omega^{*-1} \right]_C \Omega^{-1}.$$

<u>Proof</u>: Observe that

$$J(D) = E \left\{ \left| (L_3 D L_1 - L_2)\alpha + L_3 D\beta \right|^2 + E \left\{ \left| L_4 D L_1 \alpha + L_4 D\beta \right|^2 \right\} \right.$$

$$= \text{tr } X_1 + \text{tr } X_2$$

where

$$X_1 = L_3 D L_1 Q_\alpha L_1^* D^* L_3^* + L_3 D Q_\beta D^* L_3^* + L_3 D L_1 Q_{\alpha\beta} D^* L_3^* + L_3 D Q_{\beta\alpha} L_1^* D^* L_3^*$$

$$+ L_4 D L_1 Q_\alpha L_1^* D^* L_4^* + L_4 D Q_\beta D^* L_4^* + L_4 D L_1 Q_{\alpha\beta} D^* L_4^* + L_4 D Q_{\beta\alpha} L_1^* D^* L_4^*$$

and

$$X_2 = -L_3 D L_1 Q_\alpha L_2 - L_2 Q_\alpha L_1^* D^* L_3^* - L_2 Q_{\alpha\beta} D^* L_3^* - L_3 D Q_{\beta\alpha} L_2^* + L_2 Q_\alpha L_2^*.$$

Observe also that

$$\text{tr } X_1 = \text{tr } \left\{ (L_3^* L_3 + L_4^* L_4) D Q D^* \right\}$$

where

$$Q = L_1 Q_\alpha L_1^* + Q_\beta + L_1 Q_{\alpha\beta} + Q_{\beta\alpha} L_1^*.$$

Taking into account that

$$L_3^* L_3 + L_4^* L_4 = \Lambda^* \Lambda$$

it follows that $J(D)$ can be rewritten as

$$J(D) = \text{tr } \{ \Lambda D Q D^* \Lambda^* + X_2 \}.$$

From here, using some straightforward algebraic manipulations plus a completion of the square argument similar to that used in [Bo 1] one obtains

$$J(D) = J_o(D) + \text{tr } \{ \Pi_o Q \Pi_o^* + L_2 Q_\alpha L_2^* \}$$

where

$$\Pi_o = \Lambda^{*-1} L_3^* (L_2 Q_\alpha L_1^* + L_2 Q_{\alpha\beta}) Q^{-1}$$

and

$$J_o(D) = \text{tr} \{(\Lambda D - \Pi_o) Q (\Lambda D - \Pi_o)^*\}.$$

Observing now that the term $\text{tr} \{-\Pi_o Q \Pi_o + L_2 Q_\alpha L_2^*\}$ is independent of D, it follows that the problem of minimizing $J(D)$ is equivalent to that of minimizing $J_o(D)$. The noncausal solution of this problem is obviously given by $D = \Lambda^{-1} \Pi_o$. The causal solution corresponds to the solution of a best causal approximant problem with $\Phi = \Lambda D$, $Q_z = Q$, and $\Pi = \Pi_o$. This solution can then be obtained by applying Theorem 1.

The following corollary is interesting in that it provides a useful relation between the best causal approximant problem and a special version of the basic optimization problem. This result will be applied in the study of the Kalman optimal regulator problem (Theorem 5).

<u>Corollary 2</u>: If $\alpha = \beta$, $L_1 = 0$ and there exists a causal and causally invertable Λ such that

$$L_3^* L_3 + L_4^* L_4 = \Lambda^* \Lambda$$

then a sufficient condition for D_o to be a solution of the basic optimization problem is that

$$\Lambda D_o = \Phi_o$$

where Φ_o is a solution of the best causal approximant problem with $Q_z = Q_\alpha$ and $\Pi = \Lambda^{*-1} L_3^* L_2$.

Remark 1: Using lemma 7 one can see that the invertibility require-
ments on $L_3^* L_3 + L_4^* L_4$, Ω and Λ are not essential to the validity of
Theorem 2, nor to that of Corollary 2. They are merely used to sim-
plify notations. If these requirements are not satisfied one would
have to replace Λ^{-1} with Λ^{-R} (right inverse) and Λ^{*-1} with Λ^{*-L} (left
inverse).

5. OPTIMAL INPUT-OUTPUT ESTIMATION AND CONTROL

In this section we will exploit Theorems 1 and 2 to obtain the solution of three fundamental optimal estimation and control problems of the input-output type. For a physical motivation of these problems and related classical background we address the reader to [Bo 1], [Yo 1] and [Po 3].

5.1 The Wiener Filter Problem

The following statement provides a generalized version of the Wiener filter problem.

Statement 3: The Wiener filter problem (Figure 3).

Given two not necessarily causal systems L_1, L_2: $[H, P^t]$, and two random processes α, $\beta\epsilon$ $(H, \Sigma, \mathcal{P})$, determine a causal filter D_o such that, for any other causal D one has $J(D_o) \leq J(D)$, where

$$J(D) = E \{|L_2\alpha - D (L_1\alpha + \beta)|^2\}.$$

The following theorem gives a sufficient condition for the solution of the generalized Wiener filter problem.

Theorem 3: Suppose that the operator

$$Q = L_1 Q_\alpha L_3^* + Q_\beta + L_1 Q_{\alpha\beta} + Q_{\beta\alpha} L_1^*$$

is invertable. Then the noncausal solution of the Wiener filter problem is given by

$$\Pi_o = L_2 (Q_\alpha L_1^* + Q_{\alpha\beta})Q^{-1}.$$

26

A causal solution is given by

$$D_o = \left[L_2 (Q_\alpha L_1^* + Q_{\alpha\beta})\Omega^{*-1}\right]_C \Omega^{-1}$$

where Ω is causal, causally invertable and such that $Q = \Omega\Omega^*$.

The proof of Theorem 3 follows directly from Theorem 2 by recognizing that the Wiener filter problem can be viewed as a specialized version of the basic optimization problem with $L_3 = 1$ and $L_4 = 0$.

Note that just as in the classical case, ([Bo 1], p. 424), the solution procedure of the generalized Wiener filter is implemented in terms of the following physical steps:

    i) calculate a prewhitening filter which reduces the

       colored noise problem to a white noise problem; let

       this be Ω;

    ii) calculate the optimal noncausal filter associated

       with the white noise problem; this is

       $\Pi = (L_2 Q_\alpha L_1^* + Q_{\alpha\beta})\Omega^{*-1}$;

    iii) calculate the best causal approximant of Π; let it be

       $[\Pi]_C$;

    iv) construct the optimal causal filter in terms of the

       cascade composition of the inverse of the prewhitening

       filter Ω^{-1}, followed by $[\Pi]_C$; that is $D_o = [\Pi]_C \Omega^{-1}$.

Thus our abstract Hilbert space generalization of the Wiener filter enjoys the same engineering interpretation which characterizes the classical transfer function development.

An additional point of interest in regard to Theorem 3 is that by setting $L_1 = 0$ and $L_2 = 1$, it also gives an immediate solution to the following optimal estimation problem (Figure 4): Given α and β, determine a causal D_o so as to minimize $E\{|D\beta - \alpha|^2\}$. This latter problem was formulated and studied in [Sa 3]. The following corollary rectifies the incorrect solution given in that reference.

<u>Corollary 2</u>: If Q_β is invertable then a sufficient condition for D_o to be a solution of the optimal estimation problem is

$$D_o = [Q_{\alpha\beta}\Omega^{*-1}]_C \Omega^{-1},$$

where Ω is causal, causally invertable and such that $Q_\beta = \Omega\Omega^*$.

5.2 The Optimal Servo Problem

We will now consider an extended version of the optimal servo problem which has been studied among others by Newton, Gould and Kaiser [Ne 1], Chang [Ch 1], and Youla [Yo 1]. Our formulation and result are formally identical to those which can be found in this latter reference (compare our Theorem 4 with Corollary 1, in [Yo 1, part II]). A formal statement of the problem is as follows.

<u>Statement P4</u>: The optimal servo problem (Figure 5).
Given the statistically independent random processes, d, ℓ, m, and $u\varepsilon$ (H, Σ, $\mathscr{P}$), and the causal systems F, F_o, L, L_o, T, T_o, and T_s, all mapping $[H, P^t] \rightarrow [H, P^t]$. Determine a causal feedback controller N_o such that

    i) N_o is causal;

    ii) $(I + FTN_o)$ admits a causal bounded inverse;

iii) For any other causal controller N satisfying i) and ii)

one has

$$J(N_o) \leq J(N)$$

where

$$J(N) = E \{ |u - y|^2 + x^2 |y|^2$$

$$y = TN (I + FTN)^{-1} [u + L_o \ell - F_o m + (L - FT_o)d] + T_o d$$

$$r = T_s N (I + FTN)^{-1} [u + L_o \ell - F_o m + (L - FT_o)d].$$

The following lemma establishes the equivalence between the optimal servo problem and the basic optimization.

<u>Lemma 11</u>: Suppose that D_o is a solution of the basic optimization problem with

$$\beta = L_o \ell - F_o m + u + (L - FT_o)d$$

$$\alpha = u - T_o d$$

$$L_1 = 0$$

$$L_2 = 1$$

$$L_3 = T$$

$$L_4 = kT_s$$

Then a sufficient condition for a causal N_o to be a solution of the Optimal Servo Problem is that $(I + FTN_o)$ admits a causal bounded inverse and

$$(I - FTD_o)N_o = D_o.$$

__Proof:__ From Figure 5 one has

$$y = TN (I + FTN)^{-1} \left[u + L_o\ell - F_o m + (L - FT_o)d\right] + T_o d$$

$$r = T_s N (I + FTN)^{-1} \left[u + L_o\ell - F_o m + (L - FT_o)d\right]$$

Using the notation

$$\beta = u + L_o\ell - F_o m + (L - FT_o)d$$

$$\alpha = u - T_o d$$

it follows

$$e = u - y = \alpha - TN_o (I + FTN_o)^{-1}\beta$$

$$r = T_s N_o (I + FTN_o)^{-1}\beta$$

Letting now $D_o = (I + N_o FT)^{-1} N_o$ we have

$$N_o = D_o (I - FTD_o)^{-1}$$

hence

$$N_o (I + FTN_o)^{-1} = D_o$$

From here we obtain

$$J(N) = E\{|e|^2 + k^2 |r|^2\} = E\{|\alpha - TD_o|^2 + |kT_s D_o\beta|^2\} = J(D_o)$$

Suppose now that D_o is a solution of the basic optimization problem and that N_o is not a solution of the optimal servo problem. Using the above equations this would imply that there exists an $\tilde{N}$ such that

$$J(\tilde{N}) < J(N_o) = J(D_o).$$

Setting $\tilde{D} = \tilde{N} \ (I + FT\tilde{N})^{-1}$ we would then have

$$J(\tilde{D}) = J(\tilde{N}) < J(D_o)$$

which is a contradiction to the hypothesis that D_o is a solution of the basic optimization problem.

The next lemma gives a sufficient condition for the solution of the basic optimization problem which lemma 11 associates with the optimal servo problem. Its proof follows directly from Theorem 2.

<u>Lemma 12</u>: Suppose that there exists causal and causally invertable Λ and Ω such that

$$T^*T + k^2 T_s^* T_s = \Lambda^* \Lambda$$

$$L_o Q_\ell L_o^* + F_o Q_m F_o^* + Q_u + (L - FT_o)Q_d \ (L - FT_o)^* = \Omega \Omega^*$$

Then the basic optimization problem associated with the optimal servo problem has the following solution

$$D_o = \Lambda^{-1} \ [\Lambda^{*-1} T^* Q_{\alpha\beta} \Omega^{*-1}]_C \Omega^{-1}$$

where

$$Q_{\alpha\beta} = Q_u + T_o Q_d \ (L - FT_o).$$

The next theorem illustrates the conditions under which the validity of the servo problem formula proposed by Youla, Jabr and Bongiorno can be extended from the rational transfer function to the more general Hilbert space setting under consideration. The proof is an immediate consequence of lemmas 11 et 12.

Theorem 4: Let the hypothesis of lemma 12 be satisfied and suppose that $(I - TFD_o)^{-1}$ admits a causal inverse.

Then a solution of the optimal servo problem is given by

$$N_o = D_o (I - FTD_o)^{-1}$$

provided that $(I + FTN_o)$ admits a causal bounded inverse .

Remark 2: Observe that using the notation $W_o = D_o \Omega$ the solution proposed by the above theorem can be rewritten as follows

$$N_o = W_o (\Omega - FTW_o)^{-1}.$$

This solution is formally identical to that which can be found in [Yo 2]. In view of this parallelism it is in order to clarify that our approach is at one time more general and more particularized than Youla's. More general in that it treats infinite dimensional and time variant systems; more particularized in that it does not allow the open loop chain to be unstable.

5.3 The Deterministic Basic Optimization Problem

The reasoning underlining the proof of our solution of the basic optimization problem can also be applied to cases where the invertibility hypothesis of that result are not satisfied. We will show this by formulating and solving a slightly more generalized version of the basic (deterministic) optimization problem studied by Porter [Po 3]. The formulation goes as follows.

Statement P5: Porter basic deterministic optimization problem

(Figure 6).

Given three causal systems $G_i: [H, P^t] \rightarrow [H, P^t]$, $i = 1, 2, 3$, and

an element $z \epsilon H$ such that $P_o z = z$; for some $P^o = (I - P_o) \epsilon R$; determine

a causal controller, D_o, such that for any other causal D one has

$$J(D_o) \leq J(D)$$

where

$$J(D) = |(G_2 - G_1 D)z|^2 + |G_3 D z|^2.$$

    The following theorem rediscovers Porter's main result, (Theorem 2,

[Po 3]) using the basic optimization problem approach.

Theorem 5: Suppose that there exists a causal and causally invertable

Λ such that

$$G_1^* G_1 + G_3^* G_3 = \Lambda^* \Lambda$$

Then a sufficient condition for D_o to be a solution of the deterministic

basic optimization problem is

$$D_o z = \Lambda^{-1} P_o [\Lambda^{*-1} G_1^* G_2 z].$$

Proof: Observe that Porter's deterministic basic optimization problem

corresponds to the special case of our (stochastic) basic optimization

problem characterized by $L_1 = 0$ and $\alpha = \beta = z$. This is precisely the

case considered in Corollary 2. Applying this corollary and taking

into account the following correspondence of notations $L_2 = G_2$, $L_3 = G_1$,

and $L_4 = G_3$, it follows that if causal Φ_o is a solution of the best

causal approximant problem with $\Pi = \Lambda^{*-1} G_1^* G_2$, then a solution of the

deterministic basic optimization problem is given by $D_o = \Lambda^{-1} \Phi_o$.

Observe also that, given $Q_z = \Omega\Omega^*$, a sufficient condition for a causal Φ_o to be a solution of the best causal approximant problem in that $\Phi_o\Omega = \Pi\Omega$; (this follows immediately from the second equation in the proof of Theorem 1). Noting that in our case $Q_z = (\cdot, z)z$, and that $\Omega = \frac{(\cdot, z)}{|z|} z$ is such that $Q_z = \Omega\Omega^*$, it follows that a sufficient condition for a causal Φ_o to be a solution of the best causal approximant problem is

$$(\cdot, z)\Phi_o z = (\cdot, z)\Lambda^{*-1}G_1^*G_2 z.$$

In addition to this from the causality of Φ_o and the hypothesis that $z = P_o z$ one has, (lemma 1),

$$\Phi_o P_o z = P_o \Phi_o z$$

hence

$$\Phi_o z = P_o \Lambda^{*-1}G_1^*G_2 z$$

and therefore

$$\Lambda^{-1}\Phi_o z = \Lambda^{-1}P_o \left[\Lambda^{*-1}G_1^*G_2 z\right].$$

This last equation implies that a sufficient condition for a causal D_o to be a solution of the deterministic optimization problem is that

$$D_o z = \Lambda^{-1}P_o \left[\Lambda^{*-1}G_1^*G_2 z\right].$$

6. OPTIMAL STATE ESTIMATION AND CONTROL

In the same spirit of the previous section we will proceed to apply Theorems 1 and 2 to three fundamental problems of the optimal state estimation and/or control type. In contrast with the usual approach, where one looks directly for an optimal state feedback strategy, and in line with [St 1] we find it more appropriate to formulate these problems as open loop problems. It will be an interesting consequence of our results that the open loop solution can be implemented in terms of a memoryless state feedback configuration.

The physical motivation of these problems and related applications are once again well documented in the classical references [Ka 1], [Ka 2], and [An 1].

6.1 The Optimal Regulator Problem

We will start with a generalized version of the optimal regulator problem originally proposed and solved by Kalman [Ka 1]. This abstract formulation and its subsequent solution can be viewed as a natural outgrowth of similar results by Steinberger, Schumitzky and Silverman [St 1]; as such, the discussion of these authors about motivations and advantages of the Hilbert resolution space approach to the problem apply "verbatim" to the present context.

<u>Statement P6</u>: The optimal regulator problem (Figure 7).
Given two statistically independent random processes ω, $\pi \epsilon (H_1, \Sigma_1, \mathbf{P}_1)$ with $Q_\omega = I$, and three causal systems

$$B: \begin{bmatrix} H_1, & P_1^t \end{bmatrix} \rightarrow \begin{bmatrix} H_2, & P_2^t \end{bmatrix}$$

$$G: \begin{bmatrix} H_2, & P_2^t \end{bmatrix} \rightarrow \begin{bmatrix} H_2, & P_2^t \end{bmatrix}$$

$$F: \begin{bmatrix} H_2, & P_2^t \end{bmatrix} \rightarrow \begin{bmatrix} H_3, & P_3^t \end{bmatrix}$$

35

with B and F memoryless and G strictly causal such that (GB, F) is
a state trajectory realization of FGB. Determine a causal optimal
open loop regulator $D_o: [H_3, P_3^t] \to [H_1, P_1^t]$ such that for any
other causal D one has

$$J(D_o) \leq J(D)$$

where

$$J(D) = E \{|y|^2 + |u|^2\}$$

with

$$y = FGB (\omega + P^o\pi - u)$$

$$u = DP_oFGB (\omega + P^o\pi)$$

Our solution of the optimal regulator problem is formalized by
the following theorem.

Theorem 6: Principle of optimality and the Kalman regulator.
Suppose that there exists a strictly causal V_1 such that

$$I + B^*G^*F^*FGB = (I + V_1^*) (I + V_1)$$

Then a sufficient condition for D_o to be an optimal open loop regulator
is

$$D_oF = (I + M_RGB)^{-1}M_R$$

where M_R is a memoryless operator such that $V_1 = M_RGB$. The optimal
control provided by this regulator can be implemented by using the
following memoryless state feedback law

$$x = GB (\omega + P^o\pi - u)$$

$$u = M_Rx.$$

Proof: Using the following correspondence of notations

$$\alpha = \beta = P_o FGB \ (\omega + P^o \pi)$$

$$L_1 = 0$$

$$L_2 = 1$$

$$L_3 = FGBP_o$$

$$L_4 = 1$$

It is easily recognized that the optimal regulator problem is a special case of the basic optimization problem (Figure 8). Moreover, by observing that

$$I + B^* G^* F^* FGB = (I + V_1^*) \ (I + V_1)$$

implies

$$I + P_o B^* G^* F^* FGBP_o = (I + P_o V_1^*) \ (I + V_1 P_o)$$

one also recognizes that this special case is precisely the one considered in Corollary 1.

Applying that result we have the following sufficient condition for a causal D_o to be the optimal regulator

$$\Lambda D_o = (I + V_1 P_o) D_o = \Phi_o$$

where Φ_o is a solution of the best causal approximant problem with $Q_z = Q_\alpha$ and Π given by

$$\Pi = \Lambda^{*-1} L_3^* L_2 = (I + P_o V_1^*)^{-1} P_o B^* G^* F^* = P_o \ (I + V_1^*)^{-1} B^* G^* F^*$$

To characterize such a Φ_0, observe that from the statistical independence of ω and π one has

$$Q_\alpha - Q_z = P_0 FGB (Q_\omega + P^0 Q_\pi P^0) G^* F^* B^* P_0$$

From here, applying Corollary 1, one finds that a sufficient condition for Φ_0 to be a solution of the best causal approximant problem under consideration is

$$\Phi_0 FGB = [\Pi P_0 FGB]_C$$

that is

$$\Phi_0 FGB = [(I + P_0 V_1^*)^{-1} P_0 B^* G^* F^* P_0 FGB]_C$$

$$= P_0 [(I + V_1^*)^{-1} B^* G^* F^* FGB]_C.$$

Noting that by applying lemma 8i) one has

$$[\Pi P_0 FGB]_C = P_0 [(I + V_1^*)^{-1} B^* G^* F^* FGB]_C = P_0 V_1$$

where, (from lemma 8), $V_1 = M_R GB$ with M_R a well defined memoryless operator. It follows

$$[\Pi P_0 FGB]_C = M_R P_0 GB$$

hence a sufficient condition for Φ_0 is

$$[\Pi P_0 FGB]_C = \Phi_0 FGB = P_0 M_R GB$$

From these equations one obtains that a sufficient condition for D_0 to be an optimal regulator is

$$D_0 F = (I + V_1 P_0)^{-1} \Phi_0 F = (I + M_R GB)^{-1} M_R$$

This formula implies in turn that the optimal control law given by D_o, u, satisfies the following equation

$$u = (I + M_R GB)^{-1} M_R P_o GB \ (\omega_o + P^o \pi)$$

hence

$$u - M_R GBu = M_R P_o GB \ (\omega + P^o \pi)$$

From here one obtains the desired memoryless state feedback strategy, namely

$$u = M_R x$$

$$x = P_o GB \ (\omega + P^o \pi - u)$$

The physical interpretation of this strategy is illustrated by the block diagram in Figure 9.

It is interesting to observe that according to the above theorem the solution of the optimal regulator problem can be implemented either by means of an open loop configuration or by a closed loop configuration using state feedback. In general the optimal closed loop configuration is characterized by a sensitivity behavior which is better than that of the open loop configuration. To illustrate, observe that if we represent with $\gamma \epsilon H_2$ the influence of a given perturbation over the state of the optimal open loop, then $(I + M_R GB)^{-1} \gamma$ will represent the influence of that same perturbation over the state of the optimal closed loop system. Using the notation $\$ = (I + M_R GB)^{-1}$ it follows that the optimal closed loop system will have a better sensitivity behavior than the optimal open loop if and only if $||\$|| < 1$, that is if and only if $I - \$^* \$ > 0$, [Po 1]. Note that

$$I - \$^* \$ = \$^* \ \{\$^{*-1} \$^{-1} - I\} \ \$.$$

From here, using the fact that $V_1 = M_R GB$ it follows

$$I - \$^* \$ = (I + V_1^*)^{-1} \{B^* G^* F^* FGB\} (I + V_1)^{-1} > 0$$

One can then conclude that the closed loop configuration of the optimal regulator does indeed offer a better sensitivity behavior than the open loop solution.

6.2 The Optimal State Estimation Problem

The optimal state estimation problem originally proposed by Kalman [Ka 1], and Kalman and Bucy [Ka 2] is in our present context formulated as follows.

<u>Statement P7</u>: The optimal state estimation problem (Figure 10). Given four statistically independent random processes $u\epsilon(H_1, \Sigma_1, \mathcal{P}_1)$, ω, $\pi\epsilon(H_2, \Sigma_2, \mathcal{P}_2)$, $\eta\epsilon(H_3, \Sigma_3, \mathcal{P}_3)$ with $Q_\omega = I$ and $Q_\eta = I$; and three causal systems

$$B: \quad [H_1, P_1^t] \rightarrow [H_2, P_2^t]$$

$$G: \quad [H_2, P_2^t] \rightarrow [H_2, P_2^t]$$

$$F: \quad [H_2, P_2^t] \rightarrow [H_3, P_3^t]$$

with B and F memoryless and G strictly causal such that (GB, F) and $(G^* F^*, B^*)$ are state and constate trajectory realizations of FGB and $B^* G^* F^*$ respectively. Determine a causal filter $D_o = D_{o1}, D_{o2})$, with

$$D_{o1}: \quad [H_3, P_3^t] \rightarrow [H_2, P_2^t]$$

$$D_{o2}: \quad [H_1, P_1^t] \rightarrow [H_2, P_2^t]$$

such that for any other causal $D = (D_1, D_2)$ one has

$$J(D_o) \leq J(D)$$

where

$$J(D) = E\{|x - \hat{x}|^2\}$$

with

$$x = GB\ (u + \omega + P^o\pi)$$

$$\hat{x} = D_1 z + D_2 u$$

and

$$z = P_o\ [FGB\ (u + \omega + P^o\pi) + \eta].$$

The following theorem is to the effect that the optimal state estimation problem can be solved via a generalized Kalman-Bucy filter.

<u>Theorem 7</u>: Suppose that there exists a strictly causal V_2 such that

$$I + FGBB^*G^*F^* = (I + V_2)\ (I + V_2^*)$$

Then a solution to the optimal state estimation problem is given by

$$D_{o1} = GM_e\ (I + FGM_e)^{-1}$$

$$D_{o2} = (I - GM_e\ (I + FGM_e)^{-1}F)GB$$

where M_e is a memoryless operator such that $V_2 = FGM_e$. The optimal state estimation can be implemented by means of the following state feedback model

$$\hat{x} = G\ [M_e\ [z - F\hat{x}] - u]$$

$$z = FGB\ (u + \omega + P^o\pi) + \eta$$

Proof: For simplicity we will treat the case where $\pi = 0$; the proof for the case $\pi \neq 0$ is obtained by applying Corollary 1 and by using an argument identical to that developed in the proof of Theorem 5. Observe that in this case

$$J(D) = E \left\{ \left| D_1 (FGB (u + \omega) + \eta) + D_2 u - GB (u - \omega) \right|^2 \right\}$$

$$= E \left\{ \left| D_1 (FGB\omega + \eta) - GB\omega - (D_2 + (D_1 F - 1)GB)u \right|^2 \right\}$$

that is, (from the statistical independence of u, ω and η),

$$J(D) = E \left\{ \left| D_1 (FGB\omega + \eta) - GB\omega \right|^2 \right\} + E \left\{ \left| (D_2 + (D_1 F - I)GB)u \right|^2 \right\}$$

It follows

$$J(D) \leq \min_{D_1} J_1(D_1) + \min_{D_1, D_2} J_0(D_1, D_2)$$

where

$$J_1(D_1) = E \left\{ \left| D_1 (FGB\omega + \eta) - GB\omega \right|^2 \right\}$$

and

$$J_0(D_1, D_2) = E \left\{ \left| (D_2 + (D_1 F - I)GB)u \right|^2 \right\}.$$

Noting that the functional $J_0(D_1, D_2)$ is minimized by any (D_1, D_2) pair such that

$$D_2 = (I - D_1 F)GB$$

one can see that the problem of minimizing $J(D_1, D_2)$ is equivalent to that of minimizing $J_1(D_1)$. This latter problem in turn is equivalent to a basic optimization problem with $L_3 = I$ and $L_4 = 0$. This, given the hypothesis of the theorem, is precisely the problem considered in Theorem 3. Taking into account that $Q_\alpha = I$, $Q_\beta = I$, $Q_{\alpha\beta} = 0$, $Q_{\beta\alpha} = 0$,

$L_1 = FGB$ and $L_2 = GB$, and applying that result we then have

$$D_{o1} = \left[GBB^* G^* F^* (I + V_2^*)^{-1} \right]_C (I + V_2)^{-1}$$

From here, using the memorylessness of F one obtains

$$FD_{o1} = \left[FGBB^* G^* F^* (I + V_2^*)^{-1} \right]_C (I + V_2)^{-1}$$

and by observing that (see lemma 8)

$$\left[FGBB^* G^* F^* (I + V_2^*)^{-1} \right]_C = V_2$$

it follows

$$FD_{o1} = V_2 (I + V_2)^{-1}.$$

Applying lemma 9 we have $V_2^* = M_e^* G^* F^*$ and therefore $V_2 = FGM_e$. It follows

$$FD_{o1} = FGM_e (I + FGM_e)^{-1}$$

hence

$$D_{o1} = GM_e (I + FGM_e)^{-1}$$

and

$$D_{o2} = (I - GM_e (I + FGM_e)^{-1} F) GB.$$

The state feedback realization of the optimal state estimation filter is obtained by first recognizing that the optimal estimation provided by (D_{o1}, D_{o2}) is equivalent to that given by the block diagram in Figure 11. Subsequently one verifies that this block diagram is equivalent to that in Figure 12.

6.3 The Stochastic Optimal Control Problem

In what follows we revisit the optimal regulator problem by considering the case where the output of the system is corrupted by noise.

<u>Statement P8</u>: The optimal regulator problem in the presence of measurement noise (Figure 13).

Given three statistically independent random processes, ω, $\pi\epsilon(H_1, \Sigma_1, \boldsymbol{\mathcal{P}}_1)$ and $\eta\epsilon(H_3, \Sigma_3, \boldsymbol{\mathcal{P}}_3)$, with $Q_\omega = I$, and $Q_\eta = I$; and three causal systems

$$B: \; [H_1, \; P_1^t] \rightarrow [H_2, \; P_2^t]$$

$$G: \; [H_2, \; P_2^t] \rightarrow [H_2, \; P_2^t]$$

$$F: \; [H_2, \; P_2^t] \rightarrow [H_3, \; P_3^t]$$

with B and F memoryless and G strictly causal such that (GB, F) and (G^*F^*, B^*) are admissable state and costate trajectory realizations of FGB and $B^*G^*F^*$ respectively. Determine a causal regulator $D_o: \; [H_3, \; P_3^t] \rightarrow [H_1, \; P_1^t]$ such that for any other causal D one has

$$J(D_o) \leq J(D)$$

where

$$J(D) = E \; \{|u|^2 + |y|^2\}$$

with

$$u = Dz$$

$$z = P_o \; [FGB \; (\omega + P^o\pi) + \eta]$$

and

$$y = z - FGBu.$$

Our next and final theorem establishes that the stochastic optimal control problem can be solved by applying the classical separation principle: the optimal stochastic controller is given by the cascade composition of the Kalman-Bucy filter followed by the Kalman regulator.

Theorem 8: Suppose that there exists strictly causal V_1 and V_2 such that

$$I + B^* G^* F^* FGB = (I + V_1^*)(I + V_1)$$

and

$$I + FGBB^* G^* F^* = (I + V_2)(I + V_2^*)$$

Then a solution of the optimal regulator problem in the presence of noise is given by

$$D_o = (I + M_R GB)^{-1} M_R GM_e (I + FGM_e)^{-1}$$

where memoryless M_R and M_e are the solutions of the optimal regulator (Theorem 6) and optimal state estimator (Theorem 7) problems respectively. The optimal control law provided by this regulator can be implemented by means of the cascade composition of the optimal state estimator followed by the optimal regulator;

that is

$$\hat{x} = G(M_e(\hat{z} - F\hat{x}) - Bu$$

$$u = M_R \hat{x}$$

$$\hat{z} = FGB(u + \omega + P^o \pi) + \eta.$$

<u>Proof</u>: Using the following correspondence of notations

$$\alpha = P_o FGB (\omega + P^o \pi)$$

$$\beta = P_o FGB (\omega + P^o \pi) + \eta$$

$$L_1 = 0$$

$$L_2 = 1$$

$$L_3 = FGBP_o$$

$$L_4 = 1$$

one can once again see that the problem under consideration is equivalent to our basic optimization problem. For simplicity and without any loss of generality, we will again confine ourselves to the case $\pi = 0$; (as in Theorems 6 and 7, the proof for the case $\pi \neq 0$ would require a slightly more involved argument based on Corollary 1 and similar to that developed in the proof of Theorem 6).

We will then proceed by applying Theorem 2. Note that

$$I + B^* G^* F^* FGB = (I + V_1^*) (I + V_1)$$

and

$$I + FGBB^* G^* F^* = (I + V_2) (I + V_2^*)$$

imply

$$I + P_o B^* G^* F^* FGBP_o = (I + P_o V_1^*) (I + V_1 P_o)$$

and

$$I + P_o FGBB^* G^* F^* = (I + V_2 P_o) (I + P_o V_2^*)$$

Using these relations and applying Theorem 2 one obtains

$$D_o = (I + V_1 P_o)^{-1} \left[(I + P_o V_1^*)^{-1} P_o B^* G^* F^* P_o FGG^* F^* P_o (I + P_o V_2^*)^{-1} \right]_C (I + V_2 P_o)^{-1}$$

In addition to this, by virtue of lemma 9, $V_1 = M_R GB$ and $V_2 = FGM_e$, and from lemma 10

$$\left[(I + V_1^*)^{-1} B^* G^* F^* FGBB^* G^* F^* (I + V_2^*)^{-1} \right]_C = M_R GM_e .$$

It follows that the optimal regulator is given by

$$D_o = P_o (I + M_R GB)^{-1} M_R GM_e (I + FGM_e)^{-1}$$

The proof that the optimal control law can be implemented in terms of a feedback loop consisting of the optimal state estimator followed by the optimal regulator is easily obtained by verifying, by inspection, that the block diagrams in Figures 14, 15 and 16 are mutually equivalent.

<u>Remark 3</u>: The statements of the problems and the associated Theorems 6, 7 and 8 can be made slightly more general by assuming $Q_\omega = Q$ and $Q_\eta = R$ with Q and R memoryless and positive definite. To illustrate how this is done, observe for example that if in the system in Figure 7, one has $Q_\omega = Q \neq I$, and $Q_\eta = R \neq I$ then one can consider the optimal control problem for the equivalent system represented in Figure 17 where $Q_{\bar\omega} = I$, $Q_{\bar\eta} = I$ and $\bar B = Q^{-1/2} B$, $\bar G = GQ^{1/2}$, $\bar F = R^{-1/2} F$. Note that $\bar B$, $\bar G$ and $\bar F$ enjoy the same state and causality properties of the original B, G and F.

CLOSURE

Using the Hilbert resolution space valued random processes
framework developed in [Sa 3] plus recent causality and state
results [De 1], [St 1], a best causal approximant and a basic
stochastic optimization problems have been formulated and solved.
These solutions provide the key to a better understanding and a
significant generalization of a number of classical estimation and
automatic control theory results. In the context of optimal
problems of the input-output type, they yield a generalized version
of the Wiener filter (compare [Bo 1] and Theorem 3), the Youla-
Jabr-Bongiorno servo problem solution procedure (compare Theorem 4
with [Y1, part II, Corollary 1]), and Porter's solution to the
deterministic basic optimal control (compare [Po 1, Theorem 3] and
our Theorem 5). In regard to estimation and control problems of
the state space feedback type, they lead to generalized versions
of the Kalman optimal regulator (compare [Ka 1, Theorem 4] with
our Theorem 6), the Kalman-Bucy filter ([Ka 2, Theorem 1] should be
compared with our Theorem 7), and the stochastic control separation
theorem ([Kw 1, Theorem 5.3, p. 390] should be compared with our Theorem 8).

These generalizations, to the effect that most of the classical
results in optimal estimation and control theory are now extended
to multivariable, distributed parameters and time variant systems
are of course not surprising. Most of them have, in fact, in one
form or the other already appeared in the technical literature. In
particular, for example, a generalization of the Kalman-Bucy filter

48

has been given by Falb [Fa 1], using the integration theory for
Banach space valued functions of Dunford and Schwartz, by Kailath
[Kai 2] in the context of the innovations approach circle of ideas,
and by Balakrishnan [Ba 3] in the setting of infinitesimal generators
on strongly continuous groups. Similarly, a deterministic version of
the principle of optimality has been given among others by Porter [Po 1]
and in a yet more comprehensive form by Steinberger, Schumitzky and
Silverman [St 1]. A generalized version of the Wiener filter has been
given by various researchers including Kailath [Kai 2], and Balakrishnan
[Ba 1]. Finally, a generalized separation principle has been offered
among others, by Balakrishnan [Ba 1], Wohnam [Wo 1], and Curtain [Cu 1].

The main contribution of the paper is in showing that our Hilbert
resolution space approach represents a worth while addition to the
available methods: it allows to obtain all these and more results as
particular cases of one central result; it retains the same physical
and intuitive character of the classical developments by Bode and Shannon
[Bo 1] and Kalman [Ka 1]; it elucidates the connections between the plant
structural properties with respect to such concepts as state, strict
causality and causality decomposition and the structural properties of
the optimal controller and/or observer. Last but not least, it opens up
a number of new interesting questions in the realm of causal factorization,
optimal control and Riccati differential equations.

REFERENCES

A 1   Anderson, B.D.Q., Moore, J.B., _Linear Optimal Control_,
        Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Ba 1   Balakrishnan, A.V., Stochastic Filtering and Control of
        Linear Systems: A General Theory, _Control Theory of
        Systems Governed by Partial Differential Equations_,
        Academic Press, 1977.

Ba 2   Balakrishnan, A.V., _Applied Functional Analysis_, Springer-
        Verlag, N.Y., 1976.

Ba 3   Balakrishnan, A.V., Lions, State Estimation for Infinite
        Dimensional Systems", _J. Computer and Systems Sciences_,
        Vol. 1, pp. 391-403, 1967.

Bar 1   Barrett, J.F., Construction of Linear Quadratic Regulators
        Using Spectral Factorization and the Return Difference Matrix,
        Techn. Report CN/75/4, The University of Cambridge, May 1975.

Bo 1   Bode, H.W., Shannon, C.E., A Simplified Derivation of Linear
        Least Square Smoothing and Prediction Theory, _Proc. 1RE_,
        Vol. 38, pp. 417-425, Apz. 1950.

Ch 1   Chang, S.S.L., _Synthesis of Optimum Control Systems_, McGraw-
        Hill, New York, 1961.

Cu 1   Curtain, R.F., Estimation and Stochastic Control for Linear
        Infinite Dimensional Systems, _Probabilistic Analysis and
        Related Topics_, Vol. 1, (Bharucha-Reid Editor), Academic Press
        (to appear).

De 1   DeSantis, R.M., Causality Theory in Systems Analysis,
        IEEE Proc., 64 (1976), pp. 36-44.

De 2   DeSantis, R.M., Porter, W.A., On the Generalization of the
        Volterra Principle of Inversion, _J. Appl. Math. Anal_,
        48 (1974), pp. 743-748.

De 3   DeSantis, R.M., Causality Structure of Engineering Systems,
        SEL Tech, Rep. 67, The University of Michigan, Ann Arbor,
        Michigan, Sept. 72.

De 4   DeSantis, R.M., Porter, W.A., On Time Related Properties of
        Non-Linear Systems, _SIAM J. Appl. Math._, Vol. 24, pp. 188-206,
        1973.

De 5   DeSantis, R.M., On A Generalized Volterra Equation in Hilbert
        Space, _Proc. AMS_, Vol. 38, pp. 563-570, 1973.

Du 1   Duttweiler, D., Reproducing Kernel Hilbert Space Techniques
       for Detection and Estimation Problems, Ph. D. Thesis,
       Stanford Univ., Stanford, 1970.

Fa 1   Falb, P. Infinite Dimensional Filtering: the Kalman-Bucy
       Filter in Hilbert Space, Information and Control, Vol. 11,
       pp. 102-137, 1967.

Go 1   Gohberg, I.Z., Krein, M.G., Theory of Volterra Operators in
       Hilbert Space and Applications, (translation), Amer. Math.
       Soc., Vol. 18, 1969.

Ka 1   Kalman, R.E., Contributions to the Theory of Optimal Control,
       Bol. Soc. Mat. Mexicana, Vol. 5, No. 1, 1960, pp. 102-119.

Ka 2   Kalman, R.E., Bucy, R.S., New Results in Linear Filtering
       and Prediction Theory, Trans. ASME, Series D, Journal of
       Basic Engineering, Vol. 82, 1961.

Kai 1 Kailath, T., Duttweiler, D., An RKHS Approach to Detection
       and Estimation Problems - Part III: Generalized Innovations
       Representations and a Likelihood Ratio Formula, IEEE Trans.
       on Information Theory, Vol. IT-18, No. 6, November 1972,
       pp. 730-745.

Kai 2 Kailath, T., An Innovation Approach to Least-Squares Estimation,
       Parts I and II, IEEE Trans. on Automatic Control, Vol. AC-13,
       No. 6, December 1968, pp. 646-660.

Kw 1   Kwakernaak, H., Sivan, R., Linear Optimal Control Systems,
       Wiley-Interscience, 1972.

Le 1   Leake, R.J., Duality condition established in the frequency
       domain, IEEE Transactions Information Theory, (correspondence),
       Vol. IT-11, p. 461, July 1965.

Ma 1   MacFarlane, A.G.J., Return Difference Matrix Properties for
       Optimal Stationary Kalman-Bucy Filter, Proc. IEEE 1971, Vol. 118,
       No. 2, pp. 373-376.

Ne 1   Newton, G.C., Gould, L.A., Kaiser, J.F., Analytical Design of
       Linear Feedback Controls, Wiley, New York, 1957.

Po 1   Porter, W.A., A Basic Optimization Problem in Linear Systems,
       Math. Systems Theory, 5 (1971), pp. 20-44.

Po 2   Porter, W.A., Modern Foundations of Systems Theory, New York:
       MacMillan, 1966.

Po 3   Porter, W.A., Zahm, C.L., Basic Concepts in System Theory,
       SEL Tech. Rep. 33, The University of Michigan, Ann Arbor,
       Michigan, 1969.

Po 4   Porter, W.A., Some Circuit Theory Concepts Revisited
        Intern. Journal of Control, Vol. 12, pp. 433-448, 1970.

Po 5   Porter, W.A., Nonlinear Systems in Hilbert Space, Intern.
        Journal of Control, Vol. 13, pp. 593-602, 1971.

Sa 1   Saeks, R., The Factorization Problem - A Survey, IEEE Proc.,
        64 (1974), pp. 90-95.

Sa 2   Saeks, R., Resolution Space Operators and Systems, Springer-
        Verlag, N.Y., 1973.

Sa 3   Saeks, R., Reproducing Kernel Resolution Space and Its
        Applications, Journal of the Franklin Institute, Vol. 302,
        No. 4, October 1976.

Sa 4   Saeks, R., Causal Factorization, Shift Operators, and
        the Spectral Multiplicity Function, Vector and Operator
        Valued Measures, (D.H. Tucker Ed.), Academic Press, New
        York, pp. 319-335, 1973.

Sa 5   Saeks, R., Causality in Hilbert Space, SIAM Rev., Vol. 12,
        pp. 357-383, 1970.

Sa 6   Saeks, R., State in Hilbert Space, SIAM Rev., Vol. 15,
        pp. 283-308, 1973.

Sa 7   Saeks, R., Fourier Analysis in Hilbert Space, SIAM Review,
        Vol. 15, pp. 604-638, 1973.

Sc 1   Schumitzky, A., On the Equivalence Between Matrix Riccati
        Equations and Fredholm Resolvents, J. Comp. and Systems
        Sciences, 2 (1968), pp. 76-87.

Sch 1   Schnure, W.K., System Identification: A State Approach,
        SEL Tech. Rys., 80, The University of Michigan, Ann Arbor,
        Michigan, 1974.

St 1   Steinberger, M., Schumitzky, A., Silverman, L., Optimal Causal
        Feedback Control of Linear Infinite-Dimensional Systems,
        (SIAM Journal on Control: to appear).

Tu 1   Tung, L.J., Saeks, R., Wiener-Hopf Techniques in Resolution
        Space, Proc. of the 2nd OTNS International Symposium, Texas
        Tech University, August 1977.

Tu 2   Tung, L.J., Random Variables, Wiener-Hopf Filtering and
        Control Formulated in Abstract Space, Ph. D. Thesis, Texas
        Tech University, Lubbock, Texas, July 1977.

Wi 1   Wiener, N., Extrapolation, Interpolation and Smoothing of
        Stationary Time Series, MIT Press, 1966.

53

Wo 1   Wohnam, W.M., "On the Separation Theorem of Stochastic
         Control", SIAM J. Control, Vol. 6, 1968, pp. 312-326.

Yo 1   Youla, D.C., Jabr, H.A., Bongiorno, J.J., Modern Wiener-
         Hopf Design of Optimal Controllers, Parts I and II,
         IEEE Trans. on Automatic Control, Vol. AC-21, No. 1,
         pp. 3-13, No. 3, pp. 319-337, Feb. and June 1976.

FIGURES CAPTIONS

1.   The best causal approximant problem

2.   The basic optimization problem

3.   The Wiener filter problem

4.   The stochastic estimation problem

5.   The optimal servo problem

6.   Porter's basic optimization problem

7.   The optimal regulator problem

8.   Establishing the equivalence between the optimal regulator
     and the basic optimization problems

9.   Kalman optimal regulator

10. The optimal state estimation problem

11. The optimal state estimator given by the Wiener filter

12. The Kalman-Bucy filter

13. The optimal regulator problem in the noisy case

14. The optimal open loop regulator in the noisy case

15. The optimal closed loop regulator in the noisy case

16. The optimal regulator in the noisy case as given by the
     separation theorem

17. Equivalent white noise problem associated with an original
     coulored noise problem
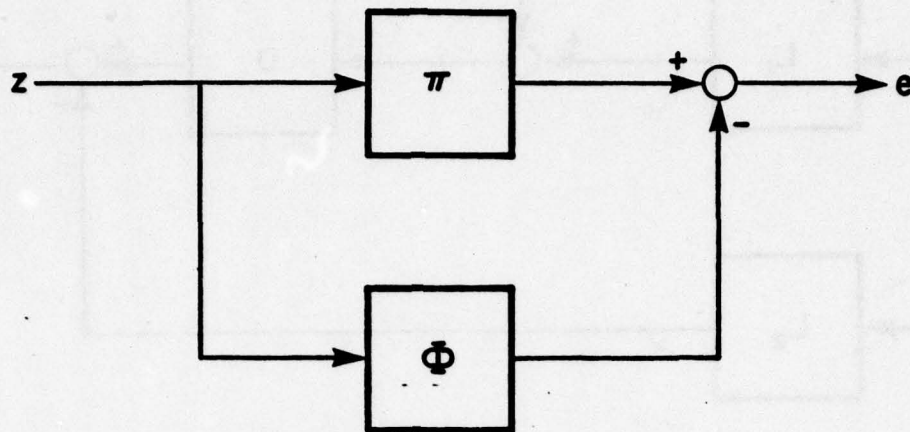
FIGURE 1: The best causal approximant problem:

given noncausal Π and the stochastic process z,

find a causal Φ_0 so as to minimize $E\{|e|^2\}$.
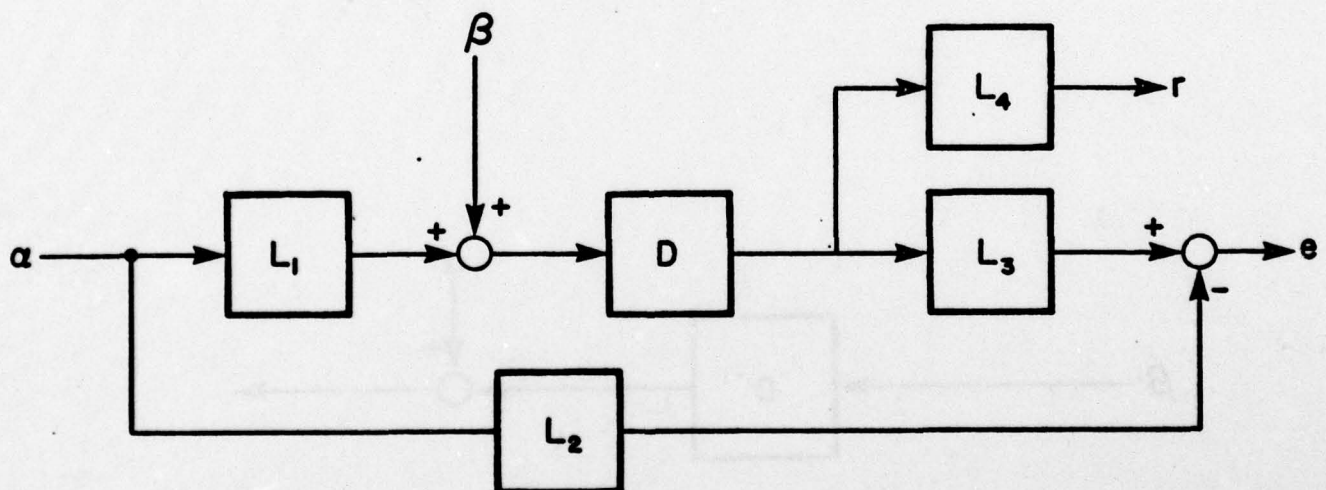
FIGURE 2: The basic optimization problem:

given L_1, α and β determine a causal D so as

to minimize $E\{|r|^2 + |e|^2\}$.
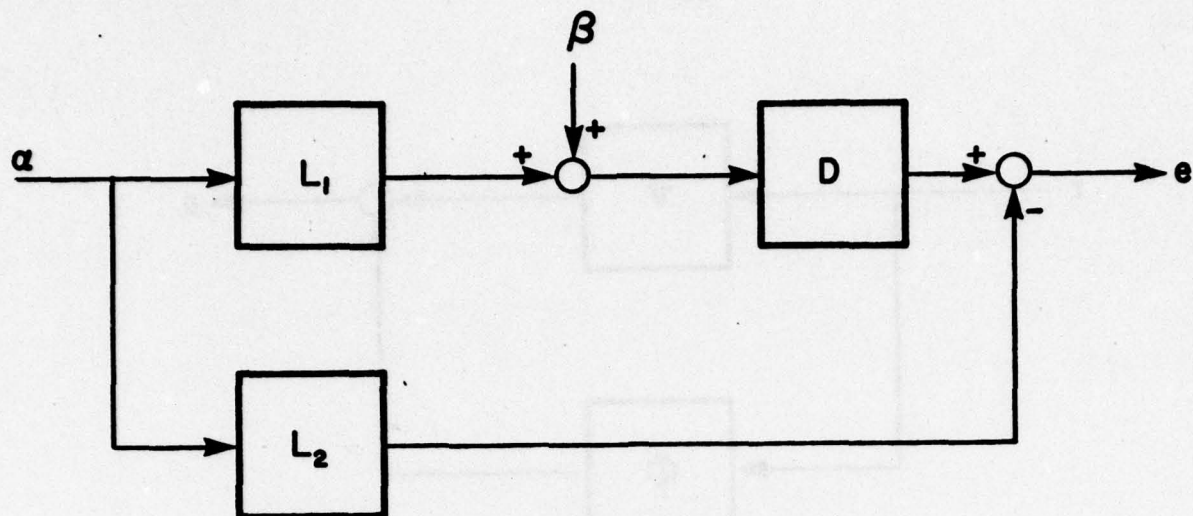
FIGURE 3: The Wiener filter problem:

given L_1, L_2, α and β determine a causal D so
as to minimize $E\{|e|^2\}$.
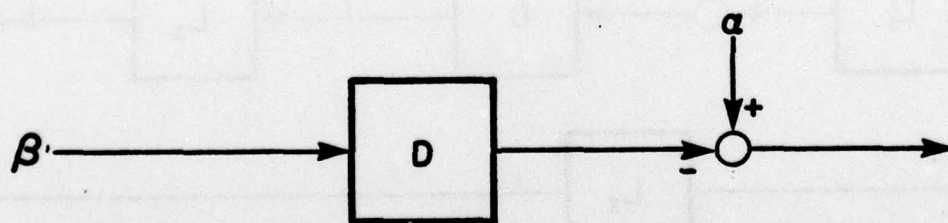
FIGURE 4: The stochastic estimation problem:

given α and β determine a causal D so as to
minimize $E\{|D\beta - \alpha|^2\}$.
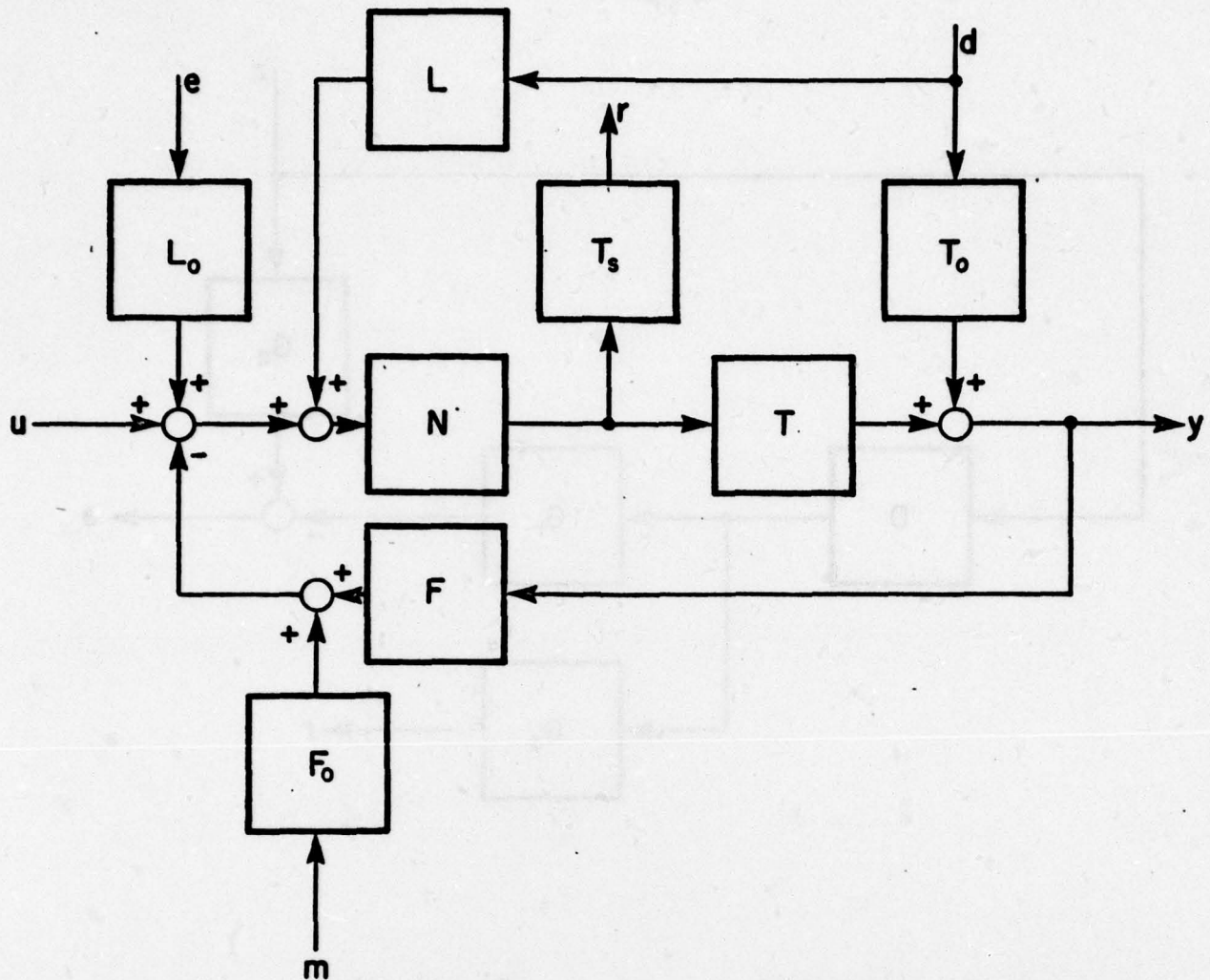
d — disturbance   T — the plant

l, m — instrumentation noise   F — transducer system

r — saturation related signal L, L_o, F_o, T_o — noise related system

u — desired output

y — actual output

FIGURE 5: The optimal servo problem:

given L_o, L, T_o, T, T_s, F, F_o and u determine a causal N

so as to minimize $E \{ |u - y|^2 + k^2 |r|^2 \}$.
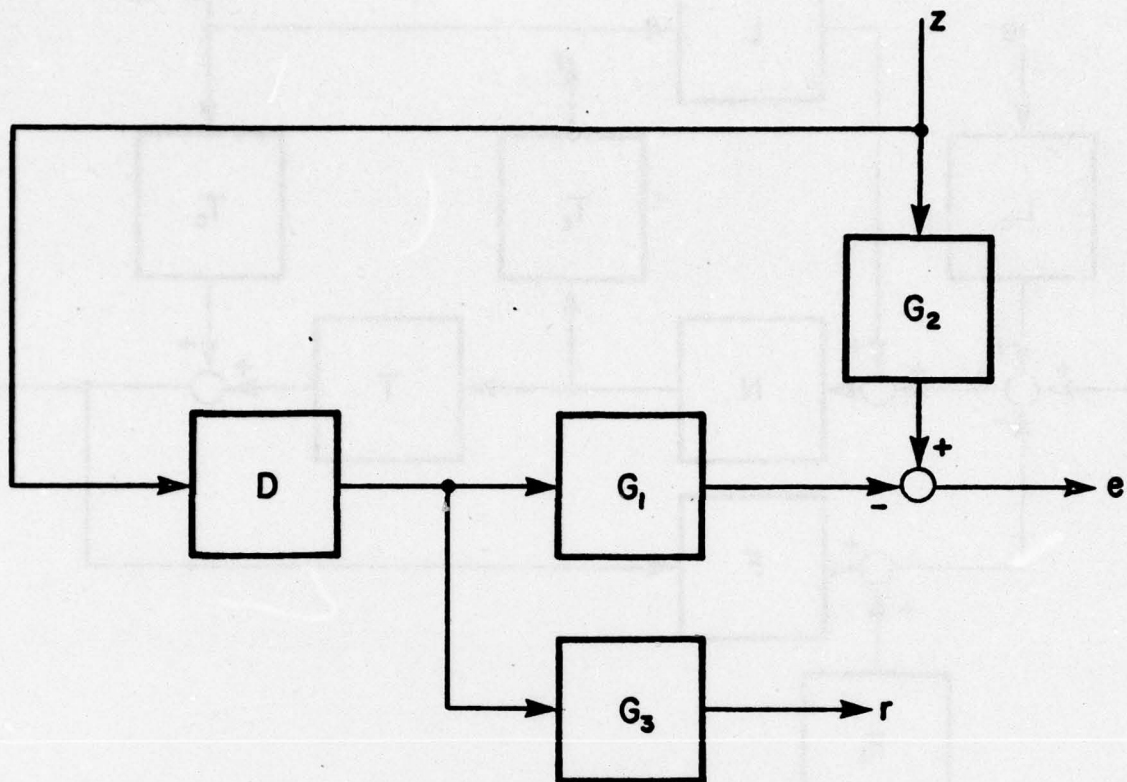
e — error signal G_1 — the plant

r — saturation related signal G_2 — the desired system

z — input

FIGURE 6: Porter's basic optimization problem:

given G_i, i — 1, 2, 3 and z determine D so as to

minimize E $\{|e|^2 + |r|^2\}$.
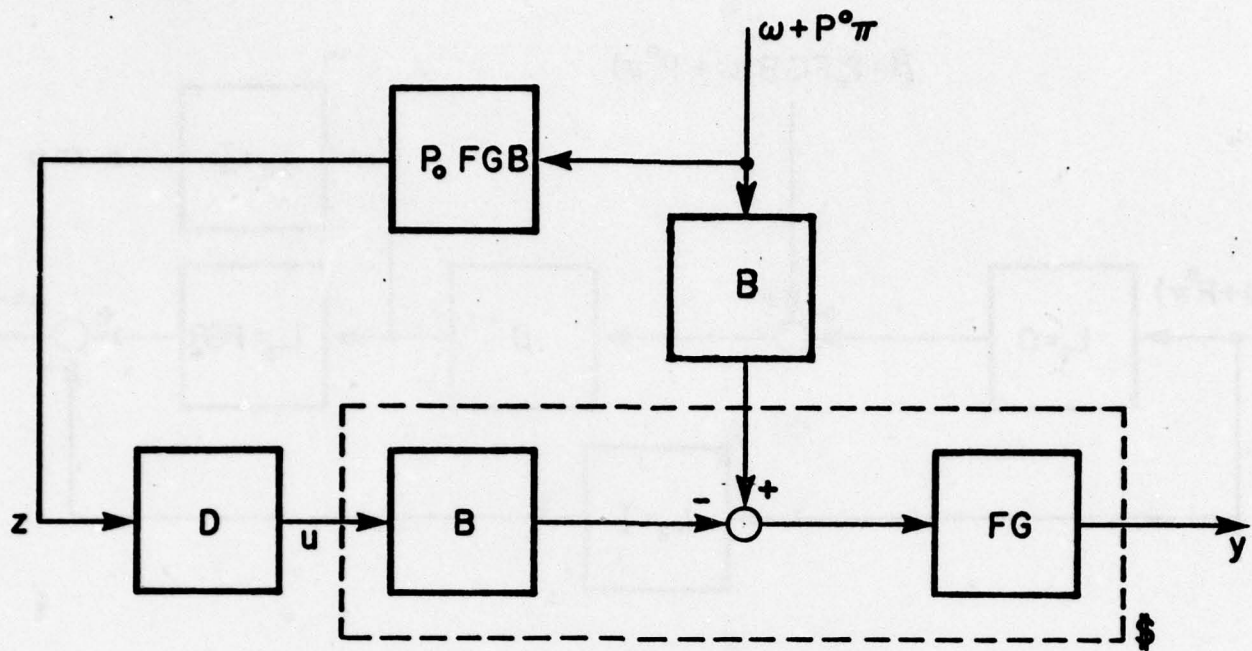
z — influence over the future output of the plant of the past

input, $(P^o\pi)$, and the perturbation, (ω); $z = P_o FGB (\omega + P^o\pi)$

y — output of the plant   D — open loop regulator

u — control $ — the plant

<u>FIGURE 7</u>: The optimal regulator problem:

knowing $ and z determine D so as to minimize

$E \{|u|^2 + |y|^2\}$
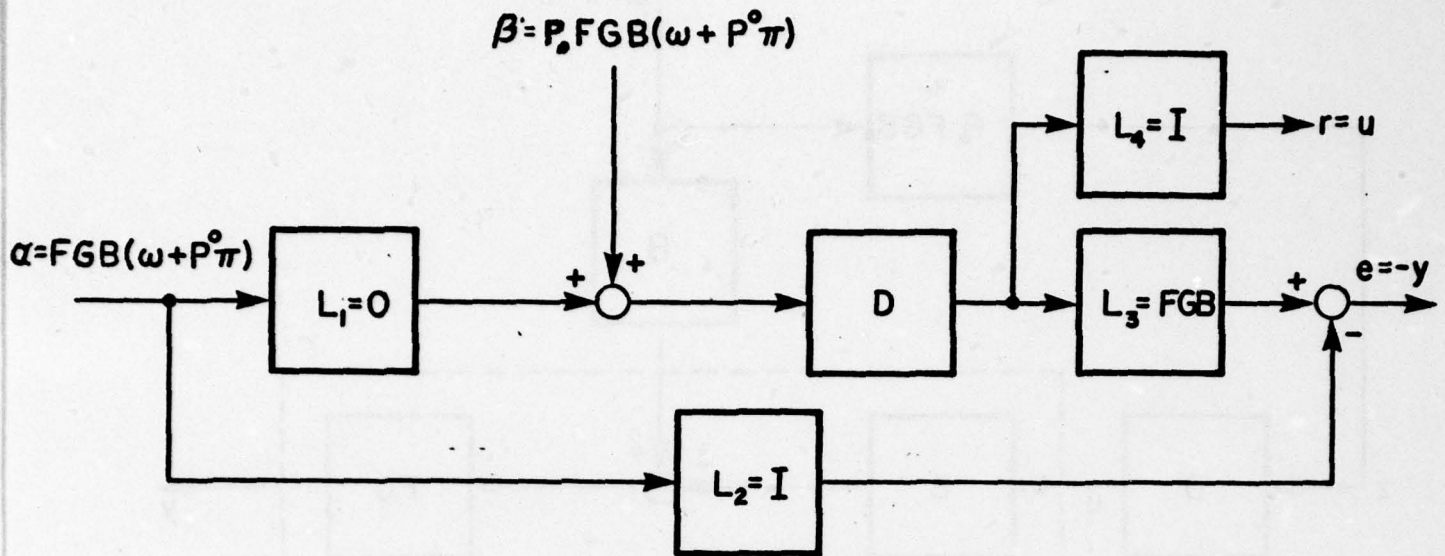
FIGURE 8: Establishing the equivalence between the optimal
regulator and the basic optimization problems.
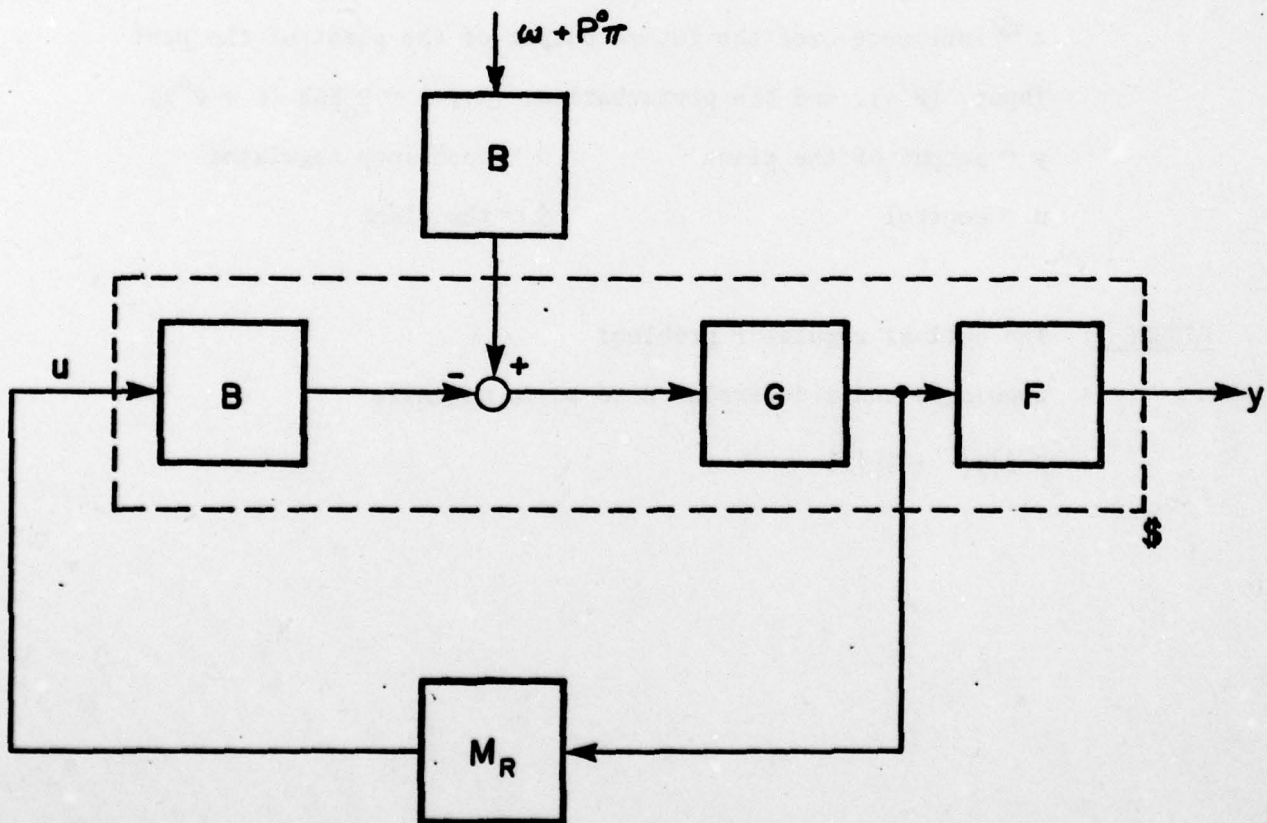
FIGURE 9: Kalman optimal regulator.
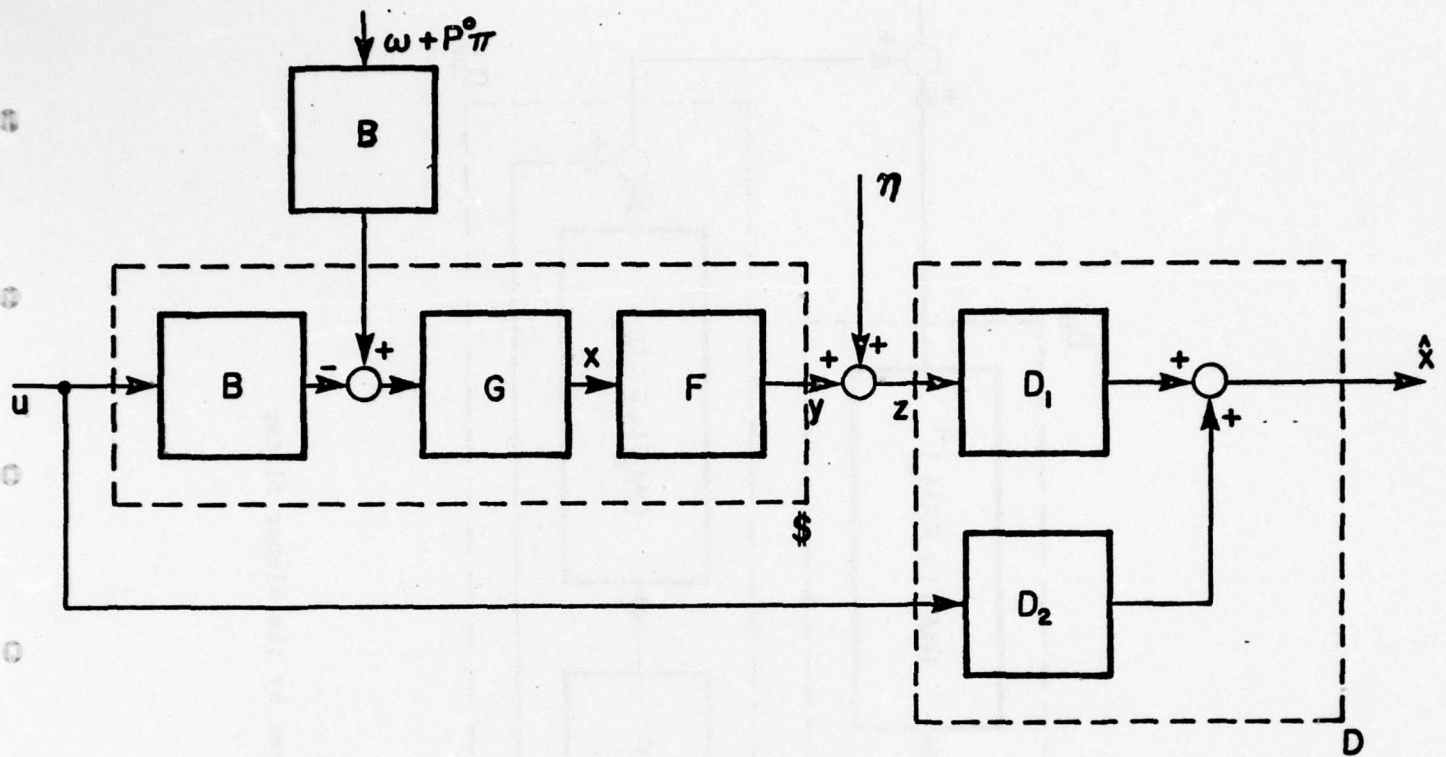
$$u = \text{plant input} \qquad\qquad z = \text{available output date}$$
$$\omega = \text{perturbation} \qquad\qquad \hat{x} = \text{estimated state}$$
$$P^o\pi = \text{past input} \qquad\qquad D = \text{state estimator}$$
$$x = \text{plant state} \qquad\qquad \$ = \text{the plant}$$
$$y = \text{plant input}$$
$$\eta = \text{measurement noise}$$

FIGURE 10: The optimal state estimation problem:

knowing $, z and u determine a causal D so as to

minimize $E\{|\hat{x} - x|^2\}$.
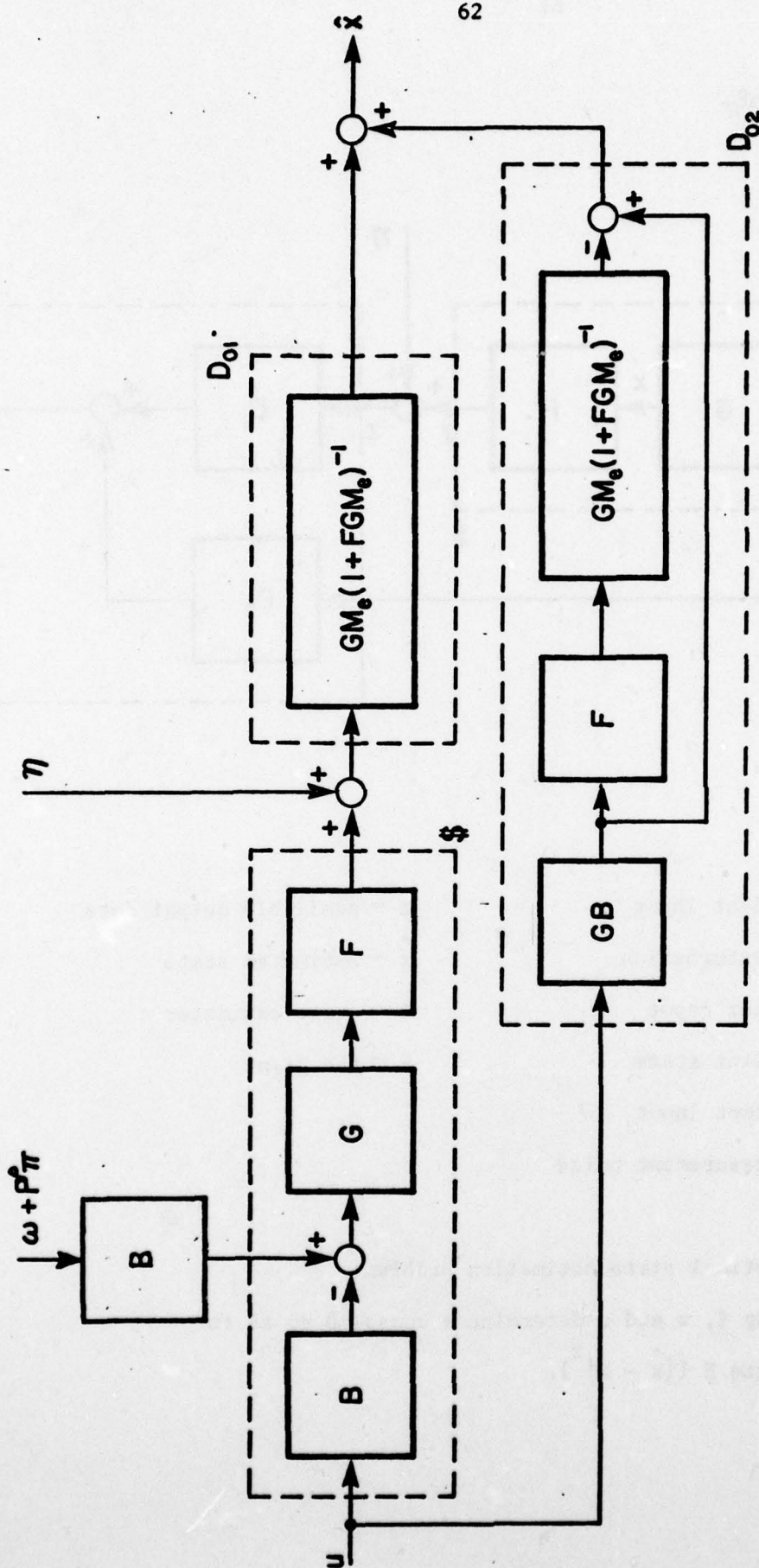
FIGURE 11: The optimal state estimator given by the Wiener filter
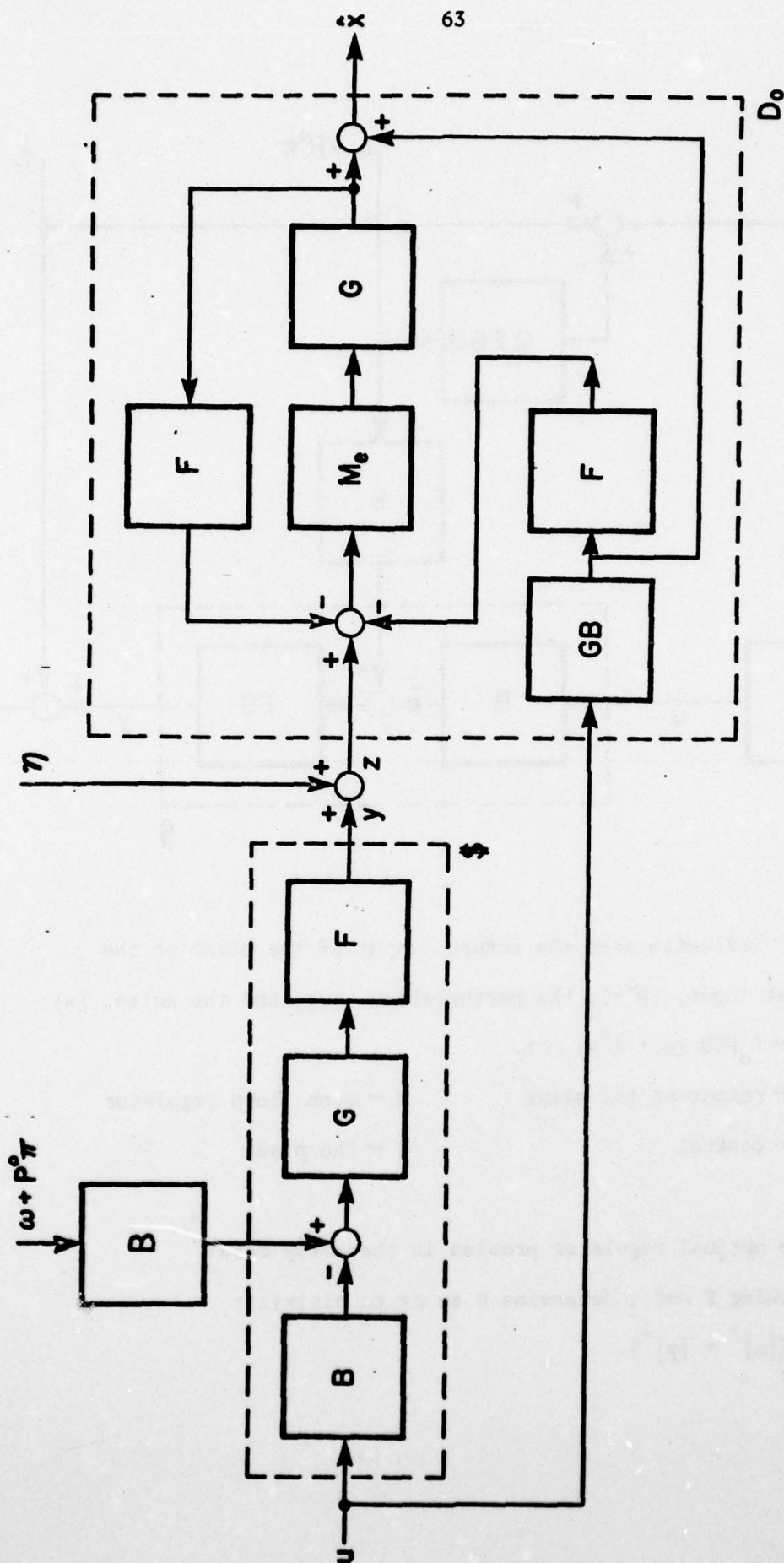
FIGURE 12: The Kalman-Bucy filter
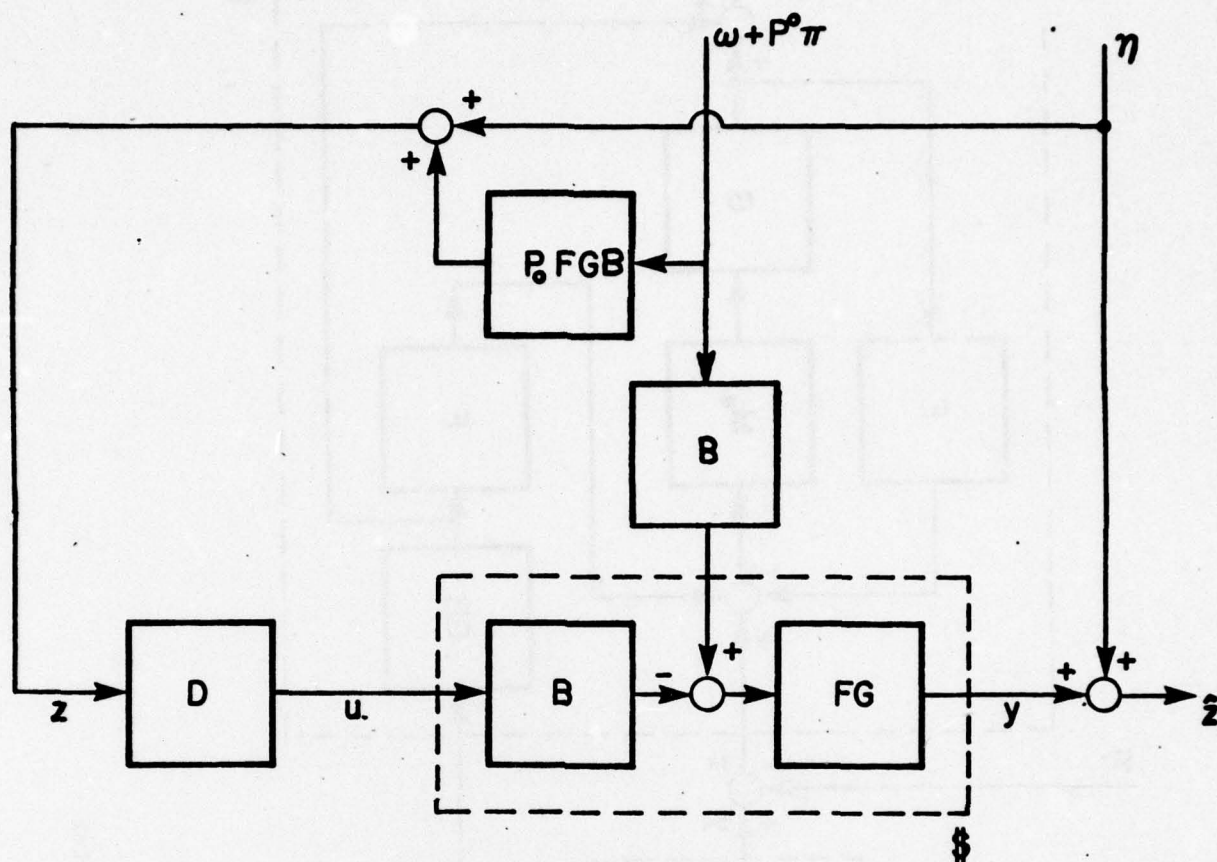
z — influence over the future output of the plant of the

past input, $(P^o\pi)$, the perturbation, (ω), and the noise, (η);

$z = P_oFGB\ (\omega + P^o\pi) + \eta$.

y — output of the plant D — open loop regulator

u — control $\$$ — the plant

FIGURE 13: The optimal regulator problem in the noisy case:
knowing $\$$ and z determine D so as to minimize
$E\ \{|u|^2 + |y|^2\}$.
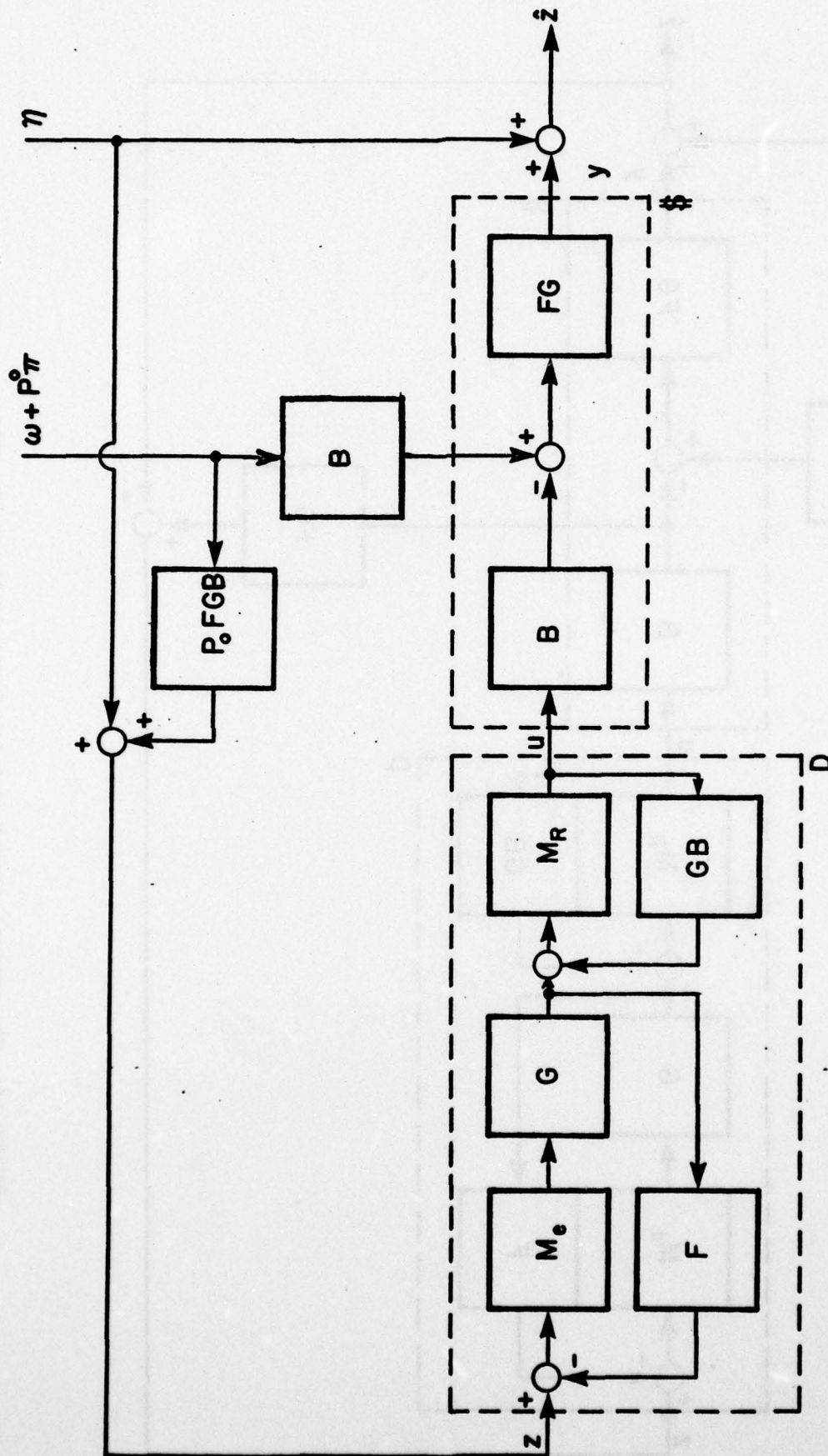
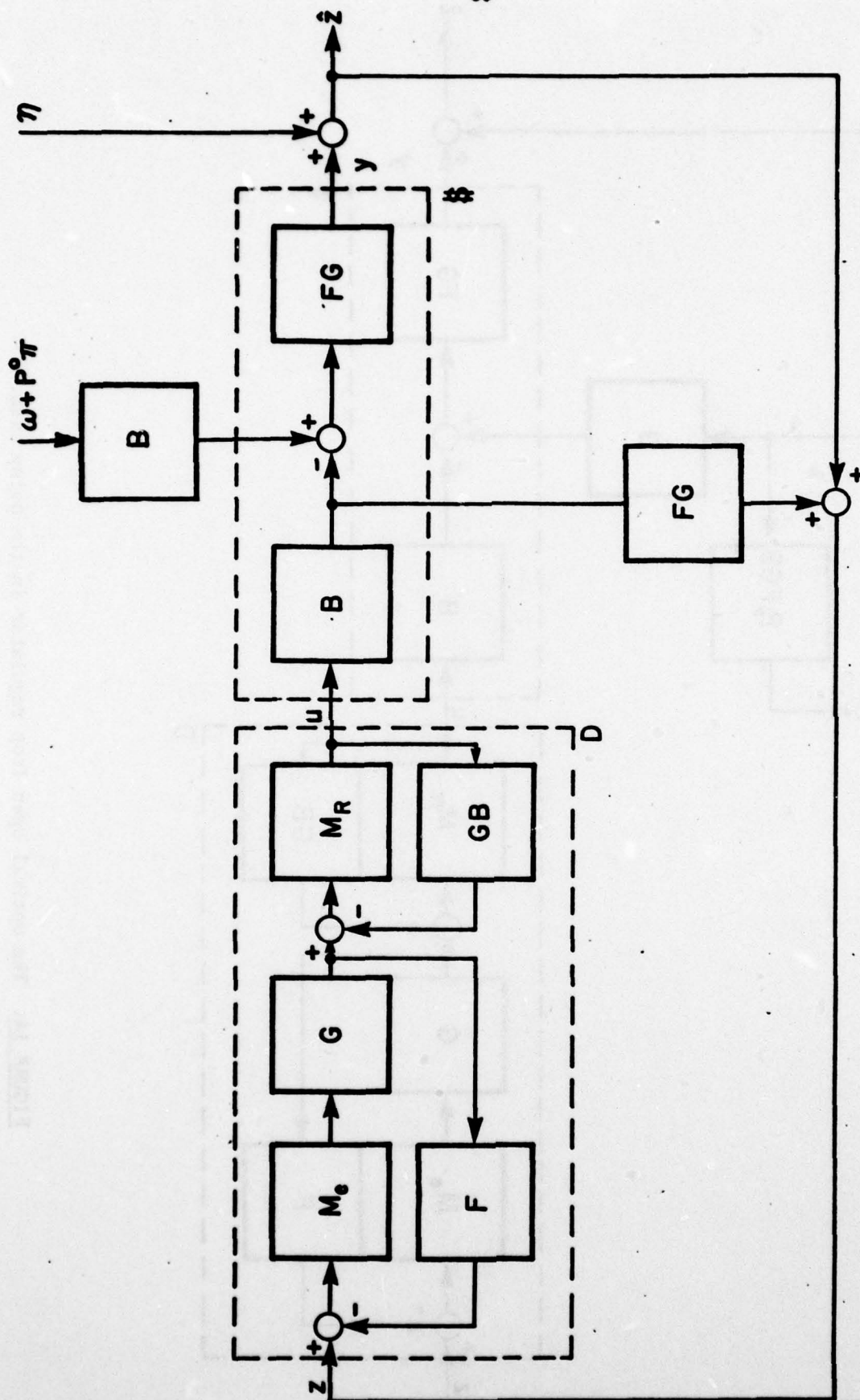FIGURE 14: The optimal open loop regulator in the noisy case.

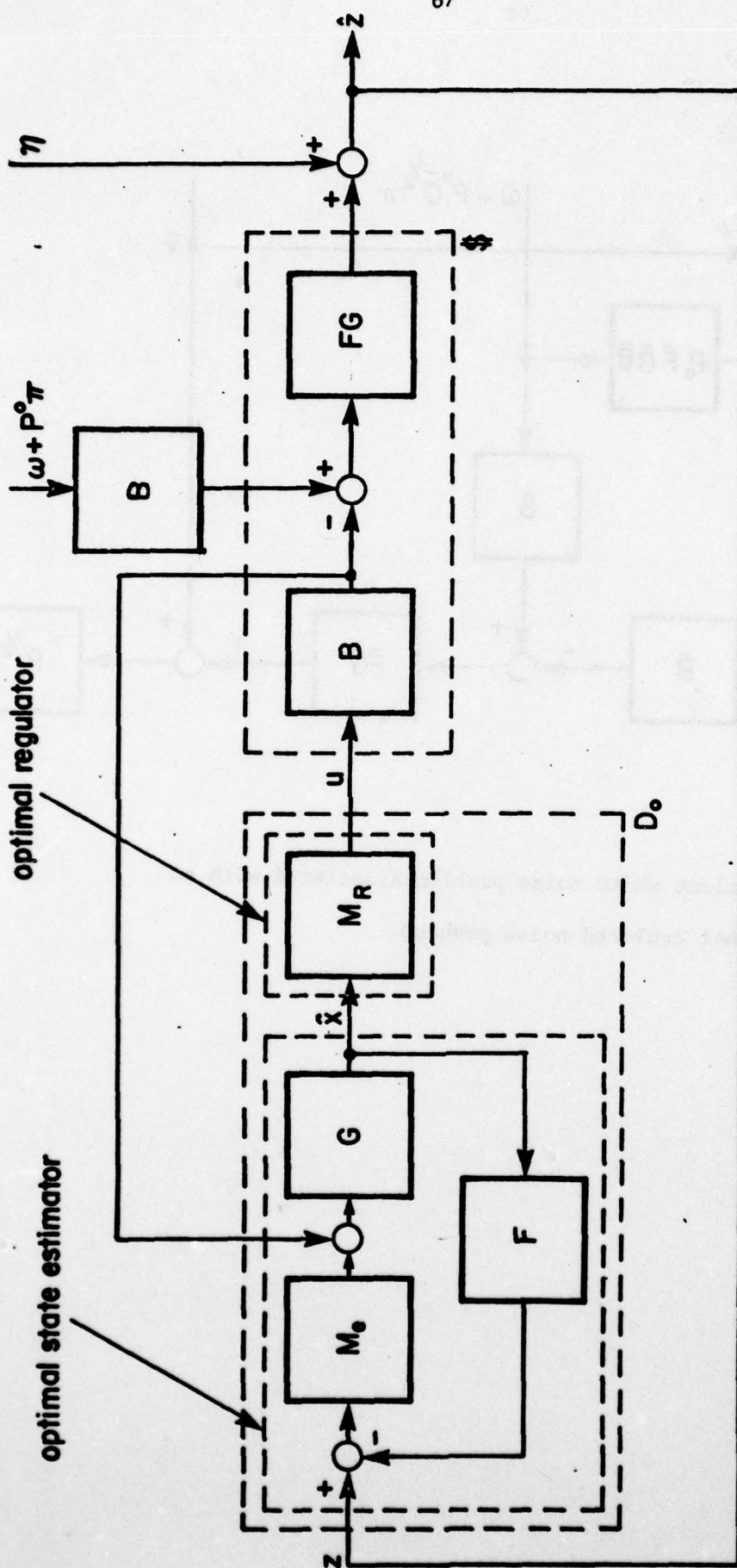FIGURE 15: The optimal closed loop regulator in the noisy case.

FIGURE 16: The optimal regulator in the noisy case as given by the separation theorem.
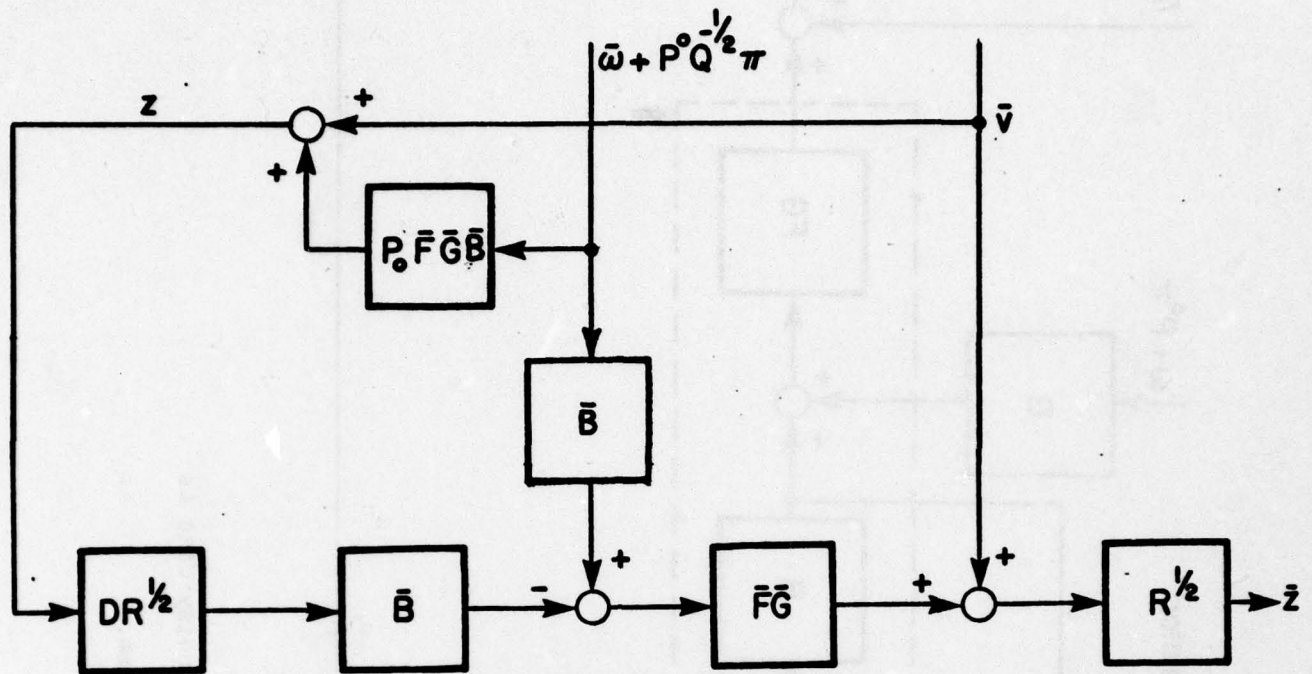
FIGURE 17: Equivalent white noise problem associated with an original coulored noise problem.